

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 22:14:19 ; Search time 24.3966 Seconds  
(without alignments)  
755.549 Million cell updates/sec

Title: US-10-790-589-2

Perfect score: 1927

Sequence: 1 MTAKPLRTVLSLFFALSGV.....KNHYIYQAFHKLRYGFKC 354

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	607.5	31.5	353	1	P56542 mus musculus
2	604.5	31.4	364	1	DRN2 MOUSE
3	596.5	31.0	360	1	O62855 sus scrofa
4	535	27.8	375	1	O00115 homo sapien
5	470.5	24.4	367	1	Q17778 caenorhabdi
6	462.5	24.0	378	1	P34387 caenorhabdi
7	443	23.0	276	1	P34508 caenorhabdi
8	101.5	5.3	475	1	P56541 bos taurus
9	94	4.9	1385	1	Q45760 bacillus th
10	94	4.9	2165	1	P28887 human respi
11	93.5	4.9	582	1	P19252 pisum sativ
12	90	4.7	1276	1	P97260 cricetulus
13	90	4.7	1445	1	P23470 homo sapien
14	88.5	4.6	1435	1	P47729 mycoplasma
15	88	4.6	3388	1	P12823 d genome po
16	87.5	4.5	175	1	P75143 mycoplasma
17	87.5	4.5	470	1	P33035 fireyella d
18	87.5	4.5	957	1	Q6R553 rattus norv
19	87.5	4.5	1278	1	Q12770 homo sapien
20	87	4.5	900	1	P47534 mycoplasma
21	87	4.5	2109	1	P16379 vesicular s
22	86.5	4.5	360	1	P50773 human papil
23	86.5	4.5	1386	1	P75613 mycoplasma
24	85	4.4	867	1	P04025 simian retr
25	85	4.4	1442	1	Q05909 mus musculu
26	84.5	4.4	541	1	Q25861 plasmodium
27	84	4.4	675	1	P51813 homo sapien
28	84	4.4	956	1	Q99j87 mus musculu
29	84	4.4	1025	1	P53708 homo sapien
30	84	4.4	5376	1	Q88799 mus musculu
31	83.5	4.3	394	1	P49150 desmodus ro
32	83.5	4.3	524	1	P47635 mycoplasma
33	83.5	4.3	737	1	Q91225 mus musculu

34	83.5	4.3	1409	1	COPI DROME
35	83.5	4.3	1436	1	WRN_XENLA
36	83	4.3	2109	1	RRPL_VSVJH
37	82.5	4.3	980	1	POLG LIV
38	82	4.3	426	1	Y958 METJA
39	82	4.3	1473	1	OVOS CHICK
40	82	4.3	2351	1	FAB_HUMAN
41	81.5	4.2	496	1	POLG LIVK
42	81.5	4.2	496	1	POLG LIVNO
43	81.5	4.2	704	1	NEUL_PIG
44	81	4.2	428	1	SYH_BUCAP
45	81	4.2	559	1	TPA_MOUSE

## ALIGNMENTS

RESULT 1					
DRN2_MOUSE					
ID	DRN2_MOUSE	STANDARD;	PRT;	353 AA.	
AC	P56542; O55053;				
DT	15-JUL-1998 (Rel. 36, Created)				
DT	15-JUL-1998 (Rel. 36, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Deoxyribonuclease II precursor (EC 3.1.22.1) (DNase II) (Acid DNase)				
DE	(Lysosomal DNase II)				
GN	DNASE2 OR DNASE2A OR DNL2.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=98382521; PubMed=9714827;				
RA	Baker K.P., Baron W.F., Henzel W.J., Spencer S.A.;				
RT	"Molecular cloning and characterization of human and murine DNase				
RT	II.";				
RL	Gene 215:281-289 (1998).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=20363736; PubMed=10903447;				
RA	Krieser R.J., Eastman A.;				
RT	"Deoxyribonuclease II: structure and chromosomal localization of the				
RT	murine gene, and comparison with the genomic structure of the human				
RT	and three C. elegans homologs.";				
RL	Gene 252:153-162 (2000).				
CC	-1- FUNCTION: Hydrolyzes DNA under acidic conditions with a preference				
CC	for double-stranded DNA. Has a possible role in apoptosis.				
CC	-1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to nucleoside 3'-				
CC	phosphates and 3'-phosphooligonucleotide end-products.				
CC	-1- SUBCELLULAR LOCATION: Lysosomal.				
CC	-1- SIMILARITY: Belongs to the DNase II family.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; AF045741; AAC35750.1; -				
DR	EMBL; AF190459; AAF20386.1; -				
DR	MGI; MGI:1329019; Dnase2a.				
DR	InterPro; IPR004947; Dnase_II.				
DR	Pfam; PF03265; DNase II; 1.				
KW	Hydrolase; Endonuclease; Nuclease; Lysosome; Zymogen; Signal;				
KW	Apoptosis.				
FT	SIGNAL	1	19		POTENTIAL.
FT	PROPEP	20	109		BY SIMILARITY.
FT	CHAIN	110	353		DEOXYRIBONUCLEASE II.
FT	CARBOHYD	71	71		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	88	88		N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 353 AA; 38810 MW; F53CDD81CF6FCF4 CRC64;  
 Query Match 31.5%; Score 607.5; DB 1; Length 353;  
 Best Local Similarity 37.4%; Pred. No. 5.5e-45;  
 Matches 130; Conservative 59; Mismatches 146; Indels 13; Gaps 6;  
 QY 6 LRTVLISLFFALSGVLGTPETISCRNEYGEAVDWFIFYLKPKRTSKAS-EBAGLQYLIDST 65  
 Db 1 MATLSLLLAALLMW-PAEALSCGDSGQDPVFFVYKLPKRTSKAS-EBAGLQYLIDST 59  
 QY 66 RQTNWNSLYLINSTSALGRTLOHLYDTHNSNDTAYLIYNDGVP--GSVNSRYOYHAK 123  
 Db 60 SDGWDQGVYINSEGAAGRSLOPLY--RKNSQLAFLLYNDOPPKSSARDSTGCHTK 117  
 QY 124 GLLVNRTQGWFLHSHVPEKPPP---VHGVEYPTSGRRYQGTGICITFGYSQFEEDFOLL 180  
 Db 118 GVLLDQGGFWLHSHVPEKPPP---VHGVEYPTSGRRYQGTGICITFGYSQFEEDFOLL 177  
 QY 181 VLQPNVYSCFIPSTFWHKLIIYMPRCANSSSIKIPVRYLAELHSAQGLNFVHFAKSSFT 240  
 Db 178 YTYPLVDYDKLEGFAQKLPDLETIKVQHLHBPWNSVILTSQAGATFOSFAKFGK 237  
 QY 241 DDITGTAQKLTHTLLAQTKQKQELPNSCLPYHYVNIKSIQV--TSKSYFSSRODH 298  
 Db 238 DDLGSGWLAEALGTLNLOVQFQNSGILPNSCNGAYQVLDVTOTGFPGRSLTFSATEDH 297  
 QY 299 SKWCVSIKGSANRWTCIGDLNRSLSHQALRGGFICTKNHYIYQAFHKL 346  
 Db 298 SKWCVAQPG---WACVGMNKNRAETHRGGTGTCVQLPFWKAFQSL 342  
 RESULT 2  
 DRN2\_PIG  
 ID DRN2\_HUMAN STANDARD; PRT; 364 AA.  
 AC O62855;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Deoxyribonuclease II precursor (EC 3.1.22.1) (DNase II) (Acid DNase)  
 DE (Lysosomal DNase II)  
 GN DNASE2 OR DNL2.  
 OS Sus scrofa (pig).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Spine;  
 RA Wang C.C., Lu S.C., Chen H.L., Liao T.H.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Liver;  
 RX MEDLINE=98321218; PubMed=9647784;  
 RA Shikawa D., Tanuma S.-I.;  
 RT "Cloning of cDNAs encoding porcine and human DNase II.";  
 RL Biochem. Biophys. Res. Commun. 247:864-869(1998).  
 CC -!- FUNCTION: Hydrolyzes DNA under acidic conditions with a preference  
 CC for double-stranded DNA. Has a possible role in apoptosis.  
 CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to nucleoside 3'-  
 CC phosphates and 3'-phosphooligonucleotide end-products.  
 CC -!- SUBCELLULAR LOCATION: Lysosomal.  
 CC -!- SIMILARITY: Belongs to the DNase II family.  
 CC -!- DATABASE: NAME=Worthington enzyme manual;  
 CC WWW="http://www.worthington-biochem.com/HDAC/".

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AJ001387; CAA04717.1; -.  
 DR EMBL; AF060221; AAC39263.1; -.  
 DR PIR; J0205; JE0205.  
 DR InterPro; IPR004947; DNase II.  
 DR Pfam; PF03265; DNase\_II; 1.  
 KW Hydrolase; Endonuclease; Nuclease; Lysosome; Zymogen; Apoptosis;  
 KW SIGNAL.  
 FT SIGNAL 1 21 POTENTIAL.  
 FT PROPEP 22 108  
 FT CHAIN 109 364  
 FT CARBOHYD 72 72 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 171 171 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 364 AA; 40183 MW; 2B690BBFE33734D CRC64;  
 Query Match 31.4%; Score 604.5; DB 1; Length 364;  
 Best Local Similarity 37.9%; Pred. No. 1e-44;  
 Matches 136; Conservative 60; Mismatches 144; Indels 19; Gaps 9;  
 QY 6 LRTVLISLFFALSGV-LGTPEISCRNEYGEAVDWFIFYLKPKRTSKAS-EBAGLQYLID 63  
 Db 1 MATLSPLLLAALLWVPGT--LTCYSGSGQDPVFFVYKLPKRTSKAS-EBAGLQYLID 58  
 QY 64 STROTWNKSLYLINSTSALGRTLOHLYDTHNSNDTAYLIYNDGVP---GSVNSRYOY 120  
 Db 59 EESGWRDAGSINSSTGALGRSLPLY---RNTSQLAFLLYNDOPPKYRGS-QHSSNRG 114  
 QY 121 HAKGLLVNRTQGWFLHSHVPEKPPP---GYEPTSGRRYQGTGICITFGYSQFEEDIF 177  
 Db 115 HTKGVLLDQGGFWLHSHVPEKPPPSSSAAAYSPPSARTYQTLICVFFLPTFLNISR 174  
 QY 178 QLLVLQPNVYSCFIPSTFWHKLIIYMPRCANSSSIKIPVRYLAELHSAQGLNFVHFAKSS 237  
 Db 175 QLTYPVMDYDKLEGDFARKFPVLEEVVGHVLEQFWNSVILTSKAGASQSFACG 234  
 QY 238 FYTDIDFTGTAQKLTHTLLAQTKQKQELPNSCLPYHYVNIKSIQVTSKS--YFSSR 295  
 Db 235 NFGDDLISGWLAEALGTLNLOVQFQNSGILPNSCNGAYQVLDVTQIAFFPGAGNFNAT 294  
 QY 296 QDHSKWCVSIKGSANRWTCIGDLNRSLSHQALRGGFICTKNHYIYQAFHKLVLRYGFCK 354  
 Db 295 EDHSKWCVA---PERPWTGVDNMRNKEHGGGTLCALPALWKAFLKPLVKAWEPCB 350  
 RESULT 3  
 DRN2\_HUMAN  
 ID DRN2\_HUMAN STANDARD; PRT; 360 AA.  
 AC O00115; O43910;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Deoxyribonuclease II precursor (EC 3.1.22.1) (DNase II) (Acid DNase)  
 DE (Lysosomal DNase II) (R31240\_2).  
 GN DNASE2 OR DNL2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98112802; PubMed=9446563;  
 RA Yasuda T., Takeshita H., Iida R., Nakajima T., Hosomi O.,  
 RT Nakashima Y., Kishi K.;  
 RL "Molecular cloning of the cDNA encoding human deoxyribonuclease II.";  
 RN J. Biol. Chem. 273:2610-2616(1998).  
 RP [2]  
 RP SEQUENCE FROM N.A.

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC ENBL; AB004574; BAA28623.1; -
CC ENBL; AF047016; AAC77366.1; -
CC ENBL; AF045937; AAC35751.1; -
CC ENBL; AF060222; AAC39852.1; -
CC ENBL; AB008564; BAB55598.1; -
CC ENBL; AB008559; BAB55598.1; JOINED.
CC ENBL; AB008560; BAB55598.1; JOINED.
CC ENBL; AB008561; BAB55598.1; JOINED.
CC ENBL; AB008562; BAB55598.1; JOINED.
CC ENBL; AB008563; BAB55598.1; JOINED.
CC ENBL; AD000092; AAB51172.1; ALT_SEQ.
CC ENBL; BC010419; AAH10419.1; -
CC PIR; JEO206; JEO206.
CC Genew; HGNC:2960; DNASE2.
CC MIM; 126350; -
CC GO; GO:0005764; C-lysosome; TAS.
CC GO; GO:0004531; F:deoxyribonuclease II activity; TAS.
CC GO; GO:0003677; F:DNA binding; TAS.
CC GO; GO:0006259; P:DNA metabolism; TAS.
CC InterPro; IPR004947; DNase II.
CC Pfam; PF03265; DNase II; 1
CC Hydrolase; Endonuclease; Nuclease; Lysosome; Zymogen; Signal;
CC Apoptosis; Glycoprotein; Polymorphism.
CC SIGNAL 1 18
CC FT PROPEP 19 106 BY SIMILARITY.
CC FT CHAIN 107 360 DEOXYRIBONUCLEASE II.
CC FT CARBOHYD 86 85 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 266 266 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT VARIANT 314 314 R -> L (in dbSNP:1061192).
CC FT FT /FTId=VAR_012044.
CC SQ SEQUENCE 360 AA; 39581 MW; DF1BBFA8A9676EA CRC64;
CC
CC Query Match 31.0%; Score 596.5; DB 1; Length 360;
CC Best Local Similarity 38.0%; Pred.No.5e-44;
CC Matches 132; Conservative 56; Mismatches 144; Indels 15; Gaps 7;
CC
CC QY 17 LSGVLGTP--EISCRNEYGEADVFFYFKLPK-RTSKASEAGLOLYLDSTFQTWNKSL 73
CC Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 6 LAALLCPAGALTCTCYGDSQPVDVFFVYKLPALRGSGEAAQRLQYKYLDESSGGWRDGR 65
CC
CC QY 74 YLINSRSLAGRLTLOHLYDHTNSTNDTAYLYND--GVGPSVNSRYQYGHAKLLVNMRT 131
CC Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 66 ALINSEPGAAGRSLOPLY--RNSQSLAFLLYNDQPPQSKADSSWRGHTKGVLLLDHD 132
CC
CC QY 132 QGFWLHSPVKPPP-----VHGYYPTSGRRYQGTGICITGYQGFERIDFOLLVLPNIYS 188
CC Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 124 GGFWLHSPVNPFPFPASSAAAYGWPHSACTYGTQLLCVSPFPAQFSKWGKQLTYTPEWYN 183
CC
CC QY 189 CIPSTPFWKLYMPRMPCANSCLKIPVRYLAELHSAQGLNFVHPAKSSFYTDITGWI 248
CC Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 184 YQLEGIFAQEFPDLENVNVKGHVSPFWNSITLTQAGAVFOSFAKFSKFGDLYSGWL 243
CC
CC QY 249 AQKLKTHLLAQTWOKKKQLPNCSLPYHYNINKSIGVTSKS--YFSSRDHSHKWCVSIK 306
CC Db 244 AALAGTNLQVFWHTKVTGILPNCSDIQWLVNQNQIAFPGPAGPSFNSTDHSHKWCVSIPK 303
CC
CC QY 307 GSNARWTCIGDLNRSLSHQALRGGGFTCTKNHYTYQAFHKLYLRYGFC 353
CC Db 304 GP---WTCVGDNNRNGEQRGGGTLCQQLPALWKAQPLVKNYQFC 347
CC
CC RESULT 4
CC YK05 CAEEL
CC Db YK05 CAEEL STANDARD: PRT: 375 AA

```

DT	01-FEB-1994 (Rel. 28, Created)
DT	28-FEB-2003 (Rel. 41, Last sequence update)
DT	28-FEB-2003 (Rel. 41, Last annotation update)
DE	Hypothetical protein F09G8.2 in chromosome III precursor.
DE	F09G8.2.
OS	Caenorhabditis elegans.
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC	Rhabditidae; Pelodierinae; Caenorhabditis.
NCBI	NCBI_TaxID=6239;
GN	[1]
EN	SEQUENCE FROM N.A.
RP	STRAIN=Bristol N2;
RC	MEDLINE=94150718; PubMed=7906398;
EX	Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA	Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA	Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA	Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA	Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA	Latrelle P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA	Parsons J., Percy C., Riken L., Roopra A., Saunders D., Showkeen R.,
RA	Sims M., Smaldon N., Smith A., Smith M., Sonnenhammer E., Staden K.,
RA	Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA	Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA	Wohlmann P.;
RT	"2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT	elegans.";
RT	Nature 368:32-38(1994).
RN	[2]
RN	REVISIONS.
RP	Waterston R.;
RA	Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RL	[3]
RP	SEQUENCE OF 26-357 FROM N.A.
RP	MEDLINE=20363736; PubMed=10903447;
EX	Krieger R.J., Pastan A.;
RA	"Deoxyribonuclease II: structure and chromosomal localization of the
RT	murine gene, and comparison with the genomic structure of the human
RT	and three C. elegans homologs.";
RT	Gene 252:155-162(2000).
RL	-1- SIMILARITY: Belongs to the Dnase II family.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
EMBL	L11247; AAA28005.2; -;
DR	EMBL; AF220526; AAF43009.1; -;
DR	PIR; S44793; S44793.
DR	Wormpep; F09G8.2; CE20662.
DR	InterPro; IPR004947; Dnase.II.
DR	Pfam; PF03265; Dnase.II; 1.
KW	Hypothetical protein; Hydrolase; Signal.
FT	SIGNAL
FT	1 18 POTENTIAL.
FT	CHAIN 19 367 HYPOTHETICAL PROTEIN F09G8.2.
FT	CARBOHYD 253 253 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT 305 305 G -> E (IN REF. 3).
FT	SEQUENCE 367 AA; 599D082D870A7555 CRC64;
QY	Query Match 24.4%; Score 470.5; DB 1; Length 367;
QY	Best Local Similarity 29.4%; Pred. No. 3.7e-33;
QY	Matches 102; Conservative 74; Mismatches 158; Indels 13; Gaps 9;
QY	12 LLPLFASGLVLTGTPETISCRNEYGAVDWFIFKYLKRTSKASEAGLQYLDSPTQTWNK 71
QY	12 SLVYLINSTRSALGRTLQHLVDYTNSTNTDTAYLYNDGVPGSVNYROY-CHAKLLIWNR 130
QY	64 G-NDINDPNVAVGATVQSVYSDAKSNK--FWFWYSDDDP--IKSADSYRGHAKGVSLFDS 118
QY	6 VLIFSIVFTTGNGKIQCCKNWRGKSVDFVYKLPKLSGAGT--SGKEFYVFAESSDWT 63







VARIANTS SLOS MET-93; LEU-326; TRP-352 AND CYS-404.  
 MEDLINE-21109432; PubMed-11175299;  
 RA Witsch-Baumgartner M., Clara E., Löffler J., Menzel H.J., Seedorf U.,  
 RA Burn J., Gillesen-Kaesbach G., Hoffmann G.F., Fitzky B.U., Mundy H.,  
 RA Clayton P., Kelley R.I., Krajewska-Walasek M., Utermann G.;  
 RT "frequency gradients of DHCR7 mutations in patients with Smith-Lemli-  
 RT Opitz syndrome in Europe: evidence for different origins of common  
 RT mutations.";  
 Eur. J. Hum. Genet. 9:45-50(2001).  
 RN [9].  
 RP VARIANT SLOS LYS-448.  
 RX MEDLINE-22829205; PubMed-1294967;  
 RA Langius F.A., Waterham H.R., Romeijn G.J., Oostheim W., de Barse M.M.,  
 RA Dorland L., Duran M., Beemer F.A., Wanders R.J., Poll-The B.T.;  
 RT "identification of three patients with a very mild form of  
 RT Smith-Lemli-Opitz syndrome.";  
 Am. J. Med. Genet. 122:24-29(2003).  
 RL CC -!- FUNCTION: Production of cholesterol by reduction of C7-C8 double  
 CC bond of 7-dehydrocholesterol (7-DHC).  
 CC CC -!- CATALYTIC ACTIVITY: Cholesterol + NADP(+) = cholesta-5,7-dien-3-  
 CC beta-ol + NADPH.  
 CC CC -!- PATHWAY: Cholesterol biosynthesis.  
 CC CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic  
 CC reticulum.  
 CC CC -!- TISSUE SPECIFICITY: Most abundant in adrenal gland, liver, testis,  
 CC and brain.  
 CC CC -!- DISEASE: Defects in DHCR7 are the cause of Smith-Lemli-Opitz  
 CC syndrome (SLOS) [MIM:270400, 268670]; also known as SLO syndrome  
 CC or RSH syndrome. SLOS is an autosomal recessive frequent inborn  
 CC disorder of sterol metabolism with characteristic congenital  
 CC malformations and dysmorphias. All patients suffer from mental  
 CC retardation. Children with SLOS have elevated serum 7-  
 CC dehydrocholesterol (7-DHC) levels and low serum cholesterol  
 CC levels. SLOS occurs in relatively high frequency: approximately 1  
 CC in 20,000 to 30,000 births in populations of northern and central  
 CC European background. Historically, a clinical distinction often  
 CC was made between classic ("type I") SLOS and the more severely  
 CC affected ("type II") patients. There is, in reality, a clinical  
 CC and biochemical continuum from mild to severe SLOS.  
 CC CC -!- SIMILARITY: Belongs to the ERG4/ERG24 family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF096305; AAC09766.1; -  
 CC EMBL; AF034544; AAC05086.1; -  
 CC EMBL; AF110060; AAC24762.1; -  
 CC EMBL; AF067127; AAC2816.1; -  
 CC EMBL; BC000054; AAH00054.1; -  
 CC EMBL; AF062481; AAC18345.1; -  
 CC GenBank; HGNC:2860; DHCR7.  
 CC MIM; 602858; -  
 CC MIM; 270400; -  
 CC MIM; 268670; -  
 CC InterPro; IPR001171; ERG4\_ERG24.  
 CC Pfam; PF01222; ERG4\_ERG24; 1.  
 CC PROSITE; PS01017; STEROL\_REDUCT 1; 1.  
 CC PROSITE; PS01018; STEROL\_REDUCT 2; 1.  
 CC Sterol biosynthesis; Cholesterol biosynthesis; Oxidoreductase; NADP;  
 KW Transmembrane; Endoplasmic reticulum; Disease mutation.  
 FT TRANSMEM 37 59 POTENTIAL.  
 FT TRANSMEM 151 173 POTENTIAL.  
 FT TRANSMEM 178 200 POTENTIAL.  
 FT TRANSMEM 264 286 POTENTIAL.  
 FT TRANSMEM 332 354 POTENTIAL.  
 FT TRANSMEM 414 436 POTENTIAL.  
 FT TRANSMEM 51 51 P -> S (in SLOS).  
 FT VARIANT /FTID=VAR\_012717.

FT	VARIANT	93	93	T -> M (in SLOS).	
FT				/FTID=VAR_012718.	
FT	VARIANT	99	99	L -> P (in SLOS).	
FT				/FTID=VAR_012719.	
FT	VARIANT	119	119	H -> L (in SLOS).	
FT				/FTID=VAR_012720.	
FT	VARIANT	157	157	L -> P (in SLOS).	
FT				/FTID=VAR_012721.	
FT	VARIANT	244	244	G -> R (in SLOS).	
FT				/FTID=VAR_012722.	
FT	VARIANT	247	247	A -> V (in SLOS).	
FT				/FTID=VAR_012723.	
FT	VARIANT	248	248	W -> C (in SLOS).	
FT				/FTID=VAR_012724.	
FT	VARIANT	289	289	T -> I (in SLOS).	
FT				/FTID=VAR_012725.	
FT	VARIANT	326	326	V -> L (in SLOS).	
FT				/FTID=VAR_012726.	
FT	VARIANT	352	352	R -> W (in SLOS).	
FT				/FTID=VAR_012727.	
FT	VARIANT	380	380	C -> S (in SLOS).	
FT				/FTID=VAR_012728.	
FT	VARIANT	404	404	R -> C (in SLOS).	
FT				/FTID=VAR_012729.	
FT	VARIANT	410	410	G -> S (in SLOS).	
FT				/FTID=VAR_012730.	
FT	VARIANT	448	448	R -> K (in SLOS; mild).	
FT				/FTID=VAR_016975.	
FT	CONFLICT	5	5	S -> L (in REF. 2).	
FT	CONFLICT	14	14	S -> A (in REF. 5).	
FT	SEQUENCE	475 AA; 54489 MW; 7D726443834C4EEB		CRC64;	
SQ					

Query Match 5.3%; Score 101.5; DB 1; Length 475;  
 Best Local Similarity 21.4%; Pred. No. 0.33;  
 Matches 100; Conservative 48; Mismatches 147; Indels 173; Gaps 23;

QY	6	LRTVLSLFFFA---	LSGVLGTPPEISCRNEVEADWF-----	IFVKLPKRSKASE	53
DB	40	LASVIFLLFADPIVYFIMACDQYSCALT-GPVVDIVTGHARLSDIMAKTPTTRKAAQ	98		
QY	54	EAGL---	QYLYLDSRTQWNKSL-----	YLINSTRGALGRTLQH	89
DB	99	LYTLWVTQVLLYTSLPDFCHKFLPGYVGGIOGAVTAGVNVKQINGLOAWLLTHLLW	158		
QY	90	LYDTNSTDNTAYLLINDGPGSVNYSQYGHAKGLVWNRTOGFWLHSPKPPVHG	149		
DB	159	FANAHLLSWFSFTIIFDNWIP-----	LLWCAN---	ILGYAVSTFAMVKGY	200
QY	150	EYPTSGRRYGOTGICITFGYSQFEIDFQLLVLPNIYSCF-----	IPSTFWKLYI	201	
DB	201	FFPTSARDCKFTG---NFFYNTMGIET-----	NPRICKWDFKLFNRPGRGIVAWTLIN	252	
QY	202	MPRMCANSSSLKIPVRYLAELHS--AQGLNFVHFSAKSFYTDIDFTGWTIAQKLKTHLLAQ	259		
DB	253	L-----SFAAK-----QRELHSHVTNAWLVNVLQ-AIYVIDFF--W-----	NE	288	
QY	260	TWOKKKQEL-----	PSNCSLP-----	Y	276
DB	289	TWYLTIDICHDFGWLWGDVWLPYLYTIQGLYLVTHVPVQLSTPHAVGVLLGLVGY	348		
QY	277	HVYNI-----	KSIGVTSKYESS--RQDHSKWCYS--IKGSANRWTC	314	
DB	349	YIFRVANHQKDLFRRTDRCGLWGKPKVIECSYTSADGQRHSHKLLVSGFWGARHFN	408		
QY	315	IGDLNRSLLHQALR-GGGFICTKNHYIQAF-----	HKLYLRYG	351	
DB	409	VGLDMSGLAYGLACGGHLLPFYIYVAILLTHRECLDRHCASKYG	456		

RESULT 9  
 CSAA\_BACUD  
 ID CSAA\_BACUD  
 AC Q45760; STANDARD; PRT; 1385 AA.

DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DE Pesticidal crystal protein cry5Aa (insecticidal delta-endotoxin  
DE Cry5Aa) (Crystalline entomocidal protoxin) (152 kDa crystal protein).  
GN CRY5A OR CRYVA(A) OR CRVA.  
OS Bacillus thuringiensis (subsp. darmstadensis).  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=132264;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NERL B-18243 / PS17;  
RA Slack A.J., Schwab G.E., Payne J.M.;  
RT "genes encoding nematode-active toxins cloned from Bacillus  
RT thuringiensis isolate PS17.";  
RL Patent number US5281530, 25-JAN-1994.  
CC -1- FUNCTION: ENDOTOXIN WITH NEMATOCIDAL ACTIVITY.  
CC -1- DEVELOPMENTAL STAGE: The crystal protein is produced during  
CC sporulation and is accumulated both as an inclusion and as part of  
CC the spore coat.  
CC -1- MISCELLANEOUS: Toxic segment of the protein is located in the N-  
CC terminus.  
CC -1- SIMILARITY: Belongs to the delta endotoxin family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; L07025; AAA67694.1; --  
DR PIR; T18213; T18213.  
DR InterPro; IPR005638; endotoxin C.  
DR InterPro; IPR005639; endotoxin N.  
DR InterPro; IPR008979; Gal bind like.  
DR Pfam; PF03944; endotoxin\_C\_1.  
DR Pfam; PF03945; endotoxin\_N\_1.  
KW Toxin; Sporulation.  
SQ SEQUENCE 1385 AA; 152439 MW; E29FF1FC799DE95 CRC64;  
  
Query Match 4.9%; Score 94; DB 1; Length 1385;  
Best Local Similarity 20.58; Pred. No. 5.6;  
Matches 76; Conservative 48; Mismatches 134; Indels 112; Gaps 16;  
  
Qy 31 EYGEADVWFIFKLPKRTSKASEBAGLYLYLSTR-----QTWKN----- 71  
Db EVQDSVETRLYGQLP-----AVDPQAGPNYVSIDSSNPIIQINMDTKTPQGASGMNTNL 454  
Qy 72 -----SLYLINSTRSALGR-----TLQHLVDTHSTNDTAYLIYNDGVP--GSVNY 115  
Db 455 MRGSVSGLSFLQDGTSLAGMGGFADTLYSLPATH-----YLSLYGTPQTSDNY 507  
Qy 116 SROYGHAKGLLVNRTQGFMLIHSPVK---PPPVHGYEYPTSGRRYQGTGICITFGYSQF 172  
Db 508 SCHVGALGV-----STPQEAFLPIIG-----QPDEQGNVSTMGF-PF 545  
Qy 173 ERIDFQLLVLPNIYSCFIPSTFHWKLIYMPRMCANSSSLKIPRYLAELHS----- 224  
Db 546 EKASYGTVVKEWMLNGANAKLSFGQSIGIPITNVTSGEYQIRCRYASNDNTNFFNVDT 605  
Qy 225 -----AQLNFWHFHAKSSFYTDITGTWIAQKLKTHLLAQ---- 259  
Db 606 GGANPIFQIINFASVTVDNNTGVQGVVYVKSATDTSNTEIPAKTINVHLTNQSSD 665  
Qy 260 TWQKKQBLPNCSLPYH-VYNIKSIGVTSKSYFSSRQDHSKWCVSIGKSANRWTCIGDL 318  
Db 666 VFLDRIFIPSLPIYHGSYNTSS-GADDVLWSSNNMY--YDIIVNGQANS-----SSI 718  
Qy 319 NRSLLQALRG 328  
Db 719 ASSMHLNKG 728

RESULT 10  
RRPL\_HRSVA STANDARD; PRT; 2165 AA.  
ID RRPL\_HRSVA  
AC P28887;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE RNA polymerase beta subunit (EC 2.7.7.48) (Large structural protein)  
DE (L protein).  
GN L.  
OS Human respiratory syncytial virus (strain A2).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11259;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=91272488; PubMed=2053282;  
RA Stec D.S., Hill M.G. III, Collins P.L.;  
RT "Sequence analysis of the polymerase L gene of human respiratory  
RT syncytial virus and predicted phylogeny of nonsegmented  
RT negative-strand viruses.";  
RL Virology 183:273-287(1991).  
CC -1- FUNCTION: Probable component of the active polymerase. It may  
CC function in mRNA synthesis, capping, methylation and poly(A)  
CC synthesis of newly synthesized viral mRNAs, RNA editing of the P  
CC gene transcript, and protein kinase activity.  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
CC {RNA} (N).  
CC -1- SIMILARITY: Belongs to the paramyxoviruses L protein family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M75730; AAA47418.1; --  
DR PIR; A40317; RRNA2.  
DR InterPro; IPR007098; RNA\_pol monom.  
DR InterPro; IPR001016; Viral RNA\_pol\_L.  
DR Pfam; PF00946; Paramyx RNA\_pol\_1.  
DR Transferase; RNA-directed RNA polymerase.  
SQ SEQUENCE 2165 AA; 250384 MW; 5A3DC8A0805305197 CRC64;  
  
Query Match 4.9%; Score 94; DB 1; Length 2165;  
Best Local Similarity 21.5%; Pred. No. 9.7;  
Matches 58; Conservative 34; Mismatches 92; Indels 86; Gaps 12;  
  
Qy 112 SVNYSROYGHAKGLLVNRTQGFMLIHSPVKFPVPHGYEYPTS-----GRYQGTGI-- 163  
Db 1333 SVNYLHR-----LTVSSRPCEF--PASIPAYRTTN-YHFDTSINRILTEKYGDLDI 1383  
Qy 164 ---CITFGYSQFEIDFQLLVLPNIYSCFIPSTFHWKLIYMPRMCANSSSLKIPRYL 219  
Db 1384 VFQNCISFGLSLMSVVB-QFTNVCPN-----RILIFK----- 1415  
Qy 220 AELHSAQGLNFVHFAKSSFYTDITGTWIAQKLKTHLLAQWQKKQBLPNCSLPHYV 279  
Db 1416 -----LNEIHLMKPPIFTGDV-----DIHKLQVIOQHMFPLDKISLTQYV- 1457  
Qy 280 NIKSIGVTSKSYFSSRQDHSKWCVSIGKS-----ANRWTCIGDL---NRSLLH 323  
Db 1458 ---ELFJUNKTLKSGSHVNSNLLIAHKLSIDYFHTNYLTSLNLAHWILLIQLMKDSKGI 1514  
Qy 324 QALRGGGFICTKXNHYIYQAFHKLRYGFC 353  
Db 1515 EKDWGEVITDHFINKLVFNAYKTYLLC 1544



of SREBPs by site-1 protease in the Golgi.  
 CC -!- SUBUNIT: In a tight complex with SREBPs.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Moves from the  
 CC endoplasmic reticulum to the Golgi in the absence of sterols.  
 CC -!- SIMILARITY: Contains 1 sterol sensor (SSD) domain.  
 CC -!- SIMILARITY: Contains 7 WD repeats.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: U67060; AAB19103.1; -.  
 CC PIR: T18526.  
 CC GO: GO:0005783; C:Endoplasmic reticulum; ISS.  
 CC GO: GO:0005794; C:Golgi apparatus; ISS.  
 CC GO: GO:0016021; C:Integral to membrane; ISS.  
 CC GO: GO:0003754; F:Chaperone activity; NAS.  
 CC GO: GO:0045541; P:negative regulation of cholesterol biosynth. . . ; ISS.  
 CC GO: GO:0045716; P:positive regulation of low-density lipoprot. . . ; ISS.  
 CC GO: GO:0006994; P:sterol depletion response, SREBP target gen. . . ; ISS.  
 CC InterPro: IPR000731; SSD 5TM.  
 CC InterPro: IPR001680; WD40.  
 CC Pfam: PF00400; WD40; 4.  
 CC SMART: SM00320; WD40; 6.  
 CC PROSITE: PS0156; SSD; 1.  
 CC PROSITE: PS00678; WD REPEATS 1; 1.  
 CC PROSITE: PS50082; WD REPEATS 2; 1.  
 CC PROSITE: PS50294; WD REPEATS REGION; 1.  
 CC Lipid metabolism; Cholesterol metabolism; Transmembrane;  
 CC Endoplasmic reticulum; Golgi stack; Repeat; WD repeat; Glycoprotein.  
 CC  
 CC DOMAIN 1 18  
 CC TRANSMEM 19 39  
 CC DOMAIN 40 279  
 CC TRANSMEM 280 300  
 CC DOMAIN 301 312  
 CC TRANSMEM 313 333  
 CC DOMAIN 334 344  
 CC TRANSMEM 345 365  
 CC DOMAIN 366 401  
 CC TRANSMEM 402 422  
 CC DOMAIN 423 423  
 CC TRANSMEM 424 444  
 CC DOMAIN 445 518  
 CC TRANSMEM 519 539  
 CC DOMAIN 540 708  
 CC TRANSMEM 709 729  
 CC DOMAIN 730 1276  
 CC TRANSMEM 284 442  
 CC REPEAT 771 811  
 CC REPEAT 771 811  
 CC REPEAT 949 999  
 CC REPEAT 1002 1039  
 CC REPEAT 1074 1111  
 CC REPEAT 1114 1152  
 CC REPEAT 1155 1192  
 CC REPEAT 1194 1232  
 CC POLY-LEU.  
 CC POLY-ARG.  
 CC N-LINKED (GLCNAC. . . )  
 CC N-LINKED (GLCNAC. . . )  
 CC N-LINKED (GLCNAC. . . )  
 CC Y->C: ABOLISHES STEROL-RESPONSE.  
 CC D->N: ABOLISHES STEROL-RESPONSE.  
 CC MUTAGEN 443 443  
 CC SEQUENCE 1276 AA; 139512 MW; A8693F7157FF5FEC CRC64;

Query Match 4.7%; Score 90; DB 1; Length 1276;  
 Best Local Similarity 17.9%; Pred. No. 11;  
 Matches 56; Conservative 50; Mismatches 102; Indels 104; Gaps 13;

29 RNEYCEAVDWFYFKLPKRTSKASEAGLQYLIDSTRQTNWNSLYLINSTRSALGRTLQ 88

75 QCEPSQPEWYV-----GAPVAYIQIIVKSSVSPWHKNLLAVDVRFLPLSRAFQ 124  
 89 HLYDTHN-----STNDTAVLIYNDGVPGSVNVSYROYGHAKGLLVNRTQGFLLIHS 139  
 125 LVEIERHNVLRSSSGTKSEEVCLQVTDLLPG-----LRK 159  
 140 VPKFPVGHVEYPTSGRRYQGTGICITFGYSQFEEIDFOLLVQPNYISCFISTHFKL 199  
 160 LRNLLEHG-----CLLSPGNQWDERFHADPDI-----GTIHQ-- 197  
 200 IYMPRCANSSSLK-----IPVRYLAELSAQGLNFHFAKSFYTTDDIFTGWLAKLKT 254  
 198 -HEPKTLQTSATLKOLLFGVPGKY-----SGVSLYTRKRTVSYTTL-----VFQYHA 245  
 255 HLLAQTWQKKQELPS-NCSLP-----YHVYNIKISIGVT-----SKSYFSSRQ 296  
 246 KFLSSLRARLMLHLPNCSLRANLVHVFKEIGIAELIPLVTYIILLPAYIYFSTRK 305  
 297 ---DHGKWCVSI 305  
 306 IDAVKSKWGLAL 317  
 RESULT 13  
 PTPG HUMAN  
 ID PTPG HUMAN STANDARD; PRT; 1445 AA.  
 AC P23470; Q15623;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Protein-tyrosine phosphatase gamma precursor (EC 3.1.3.48) (R-PTP-  
 DE gamma).  
 DE PTPRG.  
 GN Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93180796; PubMed=8382771;  
 RA Barnea G., Silvenoinen O., Shaanan B., Honegger A.M., Canoll P.D.,  
 RA D'Eustachio P., Morse B., Levy J.B., Laforgia S., Huebner K.,  
 RA Musacchio J.M., Sap J., Schlessinger J.;  
 RT "Identification of a carbonic anhydrase-like domain in the  
 RT extracellular region of RPTP gamma defines a new subfamily of  
 RT receptor tyrosine phosphatases."  
 RL Mol. Cell. Biol. 13:1497-1506(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96429999; PubMed=8833149;  
 RA Kastury K., Ohta M., Lasota J., Moir D., Dorman T., Laforgia S.,  
 RA Druck T., Huebner K.;  
 RT "Structure of the human receptor tyrosine phosphatase gamma gene  
 RT (PTPRG) and relation to the familial RCC t(3;8) chromosome  
 RT translocation."  
 RL Genomics 32:225-235(1996).  
 RN [3]  
 RP SEQUENCE OF 836-1445 FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=91006018; PubMed=2170109;  
 RA Krueger N.X., Streuli M., Saito H.;  
 RT "Structural diversity and evolution of human receptor-like protein  
 RT tyrosine phosphatases."  
 RL EMBO J. 9:3241-3252(1990).  
 RN [4]  
 RP SEQUENCE OF 874-1118 AND 1175-1409 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=90384936; PubMed=2169617;  
 RA Kaplan R., Morse B., Huebner K., Croce C., Hawk R., Ravera M.,  
 RA Ricca G., Jaye M., Schlessinger J.;  
 RT "Cloning of three human tyrosine phosphatases reveals a multigene  
 RT family of receptor-linked protein-tyrosine-phosphatases expressed in



RA Dybvig K.;  
 RT "DNA polymerase III of Mycoplasma pulmonis: isolation and  
 RT characterization of the enzyme and its structural gene, polC.";  
 RL Mol. Microbiol. 13:843-854(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=UAB CTIP;  
 RX MEDLINE=21267165; PubMed=11353084;  
 RA Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galisson F.,  
 RA Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C.,  
 RA Blanchard A.;  
 RT "The complete genome sequence of the murine respiratory pathogen  
 RT Mycoplasma pulmonis";  
 RL Nucleic Acids Res. 29:2145-2153(2001).  
 CC -!- FUNCTION: Required for replicative DNA synthesis. This DNA  
 CC polymerase also exhibits 3' to 5' exonuclease activity.  
 CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate  
 CC + {DNA}(N).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Belongs to the DNA polymerase type-C family. PolC  
 CC subfamily.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch)).  
 CC  
 DR EMBL; U06833; AAA18178.1; --  
 DR EMBL; AL445565; CAC13848.1; --  
 DR PIR; C90596; C90596.  
 DR PIR; S54697; S54697.  
 DR MYPULIST; MYPUL 6750; --  
 DR HAMAP; MF 00356; --  
 DR InterPro; IPR006054; DnaQ.  
 DR InterPro; IPR006055; Exonuclease.  
 DR InterPro; IPR003141; PIP N.  
 DR InterPro; IPR006308; PolC\_gran\_pos.  
 DR Pfam; PF00929; Exonuclease; 1.  
 DR Pfam; PF02231; PIP N; 1.  
 DR SMART; SM00479; EXOIII; 1.  
 DR SMART; SM00481; POLIIIAC; 1.  
 DR TIGRFAMS; TIGR00573; dnaq; 1.  
 DR TIGRFAMS; TIGR01405; polC\_Gram\_pos; 1.  
 DR TransFam; TFS001405; polC\_Gram\_pos.  
 KW Nuclease; Exonuclease; Complete proteome.  
 FT DOMAIN 404 570 EXONUCLEASE.  
 FT CONFLICT 786 787 NP -> S (IN REF. 1).  
 FT CONFLICT 1233 1233 R -> KE (IN REF. 1).  
 SQ SEQUENCE 1435 AA; 166374 MW; 9E6527A88725503F CRC64;  
 Query Match 4.6%; Score 88.5; DB 1; Length 1435;  
 Best Local Similarity 22.8%; Pred. No. 17;  
 Matches 75; Conservative 42; Mismatches 121; Indels 91; Gaps 22;  
 18 SGVLGTPIS---CRNEYGEAVDWFYKLPKRTSKASEEAGLQVLYLDSTQVW-NKSL 73  
 1134 TGALGIPEFGTGFVRQMLSEA-----KPKTFADLVISIGLSH-----GTDVWINNAH 1180  
 74 YLINSRSLAGRTLOHLYDTHNSTNDTAVLIYNDGVPGSVNYS--ROYGHAKGLLVNKR 131  
 1181 YIIQS-----LGKTLQVLI---SCRDDIMVDLKKGVFDLSTIMEQVKGKLSLEQKR 1233  
 132 Q-----GF--WLHSHVFK-----PPVHGVEYPTSGR-----YGTGICITFGYSQFEEI 175  
 1234 KLEHGIENWYTESMEKIKYMPKKAHATAYVLMAWVAFYKVYVPLEYVATYFSTRTEFF 1293  
 176 DFQL-----LVLPNIYSCFISTFWKLIYMPRCANSSSL-----KIPVRYLAELHS 224  
 1294 DIEIMSKDKLTLESKIK-----ELAYRENIL-RNDNQLTTEKNTLPTLYIANEMK 1342

QY 225 AQLNFVHF-AKSFYTDIFTGWIQAQKLKTHLLAQTWQKKQELPSNCLPWHVYNIKS 283  
 Db 1343 ARGFNQINIKISLAND-----WIIDK-----NSKSLIP-----PENV--IDG 1379  
 QY 284 IGVY-SKSYFSSRQDHQKWCVCISKGSANR 311  
 Db 1380 LGETLQAKIVDSRNE--KEFLSVDFINR 1406  
 RESULT 15  
 POLG DEN2P STANDARD; PRT; 3388 AA.  
 ID AC P12823; Q88646; Q88647; Q88648; Q88650; Q88651; Q88652;  
 AC Q88653; Q88654; Q88655;  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Genome polypeptide [Contains: Capsid protein C (Core protein); Matrix  
 DE protein (Envelope glycoprotein M); Major envelope protein E;  
 DE Nonstructural proteins NS1, NS2A, NS2B, NS4A and NS4B;  
 DE Protease/helicase (EC 3.4.21.98) (NS3); RNA-directed RNA polymerase  
 DE (EC 2.7.7.48) (NS5)].  
 OS Dengue virus type 2 (strain PR159/SI).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Flavivirus.  
 CX NCBI\_TaxID=11066;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88101365; PubMed=2827375;  
 RA Hahn Y.S., Galler R., Hunkapiller T., Dalrymple J.M., Strauss J.H.,  
 RA Strauss E.G.;  
 RT "Nucleotide sequence of dengue 2 RNA and comparison of the encoded  
 RT proteins with those of other flaviviruses.";  
 RL Virology 162:167-180(1988).  
 RN [2]  
 RP C-TERMINUS OF M; E AND NS1.  
 RC STRAIN=New-Guinea;  
 RX MEDLINE=89299482; PubMed=2741348;  
 RA Wright P.J., Cauchi M.R., Ng M.L.;  
 RT "Definition of the carboxy termini of the three glycoproteins  
 RT specified by dengue virus type 2.";  
 RL Virology 171:61-67(1989).  
 CC -!- FUNCTION: The small proteins NS2A, NS2B, NS4A and NS4B are  
 CC hydrophobic, suggesting a possible membrane-related function.  
 CC NS3 and NS5 may play a role in the viral RNA replication.  
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral  
 CC precursor polypeptide, commonly with Asp or Glu in the P6  
 CC position, Cys or Thr in P1 and Ser or Ala in P1',  
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC {RNA}(N).  
 CC -!- SUBUNIT: The virion of this virus is a nucleocapsid covered by a  
 CC lipoprotein envelope. The envelope consists of two proteins:  
 CC protein M and glycoprotein E. The nucleocapsid is a complex of  
 CC protein C and mRNA.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch)).  
 CC  
 CC EMBL; M19197; AAA42962.1; --  
 DR MEROPS; S07.002; --  
 DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR001122; Flavi\_capsidC.  
 DR InterPro; IPR000336; Flavi\_Glycoprote.  
 DR InterPro; IPR000069; Flavi\_M.  
 DR InterPro; IPR001157; Flavi\_NS1.  
 DR InterPro; IPR000752; Flavi\_NS2A.  
 DR InterPro; IPR000487; Flavi\_NS2B.



DR InterPro; IPR000404; Flavi NS4A.  
DR InterPro; IPR001528; Flavi NS4B.  
DR InterPro; IPR000208; Flavi NS5.  
DR InterPro; IPR002535; Flavi propep.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR001850; Peptidase S7.  
DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro; IPR007094; RNA\_pol\_PSVir.  
DR InterPro; IPR002877; RmJ FtsJ.  
DR Pfam; PF01003; Flavi capsid; 1.  
DR Pfam; PF02832; Flavi\_glycop\_C; 1.  
DR Pfam; PF00889; Flavi\_glycop\_C; 1.  
DR Pfam; PF00949; Flavi\_glycop\_C; 1.  
DR Pfam; PF01004; Flavi M; 1.  
DR Pfam; PF00948; Flavi NS1; 1.  
DR Pfam; PF01005; Flavi NS2A; 1.  
DR Pfam; PF01002; Flavi NS2B; 1.  
DR Pfam; PF01350; Flavi NS4A; 1.  
DR Pfam; PF01349; Flavi NS4B; 1.  
DR Pfam; PF00972; Flavi NS5; 1.  
DR Pfam; PF01570; Flavi propep; 1.  
DR Pfam; PF01728; FtsJ; 1.  
DR ProDom; PD001556; Flavi\_glycop\_C; 1.  
DR ProDom; PD001496; Flavi NS1; 1.  
DR ProSITE; PS00690; DEAH ATP HELICASE; FALSE NEG.  
KW Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;  
KW Core protein; Coat protein; Envelope protein; Hydrolase; Helicase;  
KW ATP-binding; Transmembrane; Nonstructural protein.  
FT CHAIN 1 114 CAPSID PROTEIN C.  
FT PROPEP 115 205  
FT CHAIN 206 280 ENVELOPE GLYCOPROTEIN M.  
FT CHAIN 281 775 MAJOR ENVELOPE PROTEIN E.  
FT CHAIN 776 1127 NONSTRUCTURAL PROTEIN NS1.  
FT CHAIN 1128 1345 NONSTRUCTURAL PROTEIN NS2A.  
FT CHAIN 1346 1475 NONSTRUCTURAL PROTEIN NS2B.  
FT CHAIN 1476 2090 PROTEASE/HELICASE (NS3).  
FT CHAIN 2091 2376 NONSTRUCTURAL PROTEIN NS4A.  
FT CHAIN 2377 2488 NONSTRUCTURAL PROTEIN NS4B.  
FT CHAIN 2489 3388 RNA-DIRECTED RNA POLYMERASE (NS5).  
FT NP\_BIND 1668 1675 ATP (POTENTIAL).  
FT SITE 1759 1762 DEAH BOX.  
FT TRANSMEM 43 59 POTENTIAL.  
FT TRANSMEM 101 117 POTENTIAL.  
FT TRANSMEM 268 284 POTENTIAL.  
FT TRANSMEM 727 743 POTENTIAL.  
FT TRANSMEM 757 773 POTENTIAL.  
FT DISULFID 283 310 BY SIMILARITY.  
FT DISULFID 340 396 BY SIMILARITY.  
FT DISULFID 354 385 BY SIMILARITY.  
FT DISULFID 372 401 BY SIMILARITY.  
FT DISULFID 465 565 BY SIMILARITY.  
FT DISULFID 582 613 BY SIMILARITY.  
FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 347 347 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 433 433 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 905 905 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 982 982 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1134 1134 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1174 1174 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 2298 2298 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 2302 2302 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 2454 2454 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 2482 2482 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 3388 AA; 379213 MW; 79B7C87BE64D2D8F CRC64;

Query Match 4.6%; Score 88; DB 1; Length 3388;  
Best Local Similarity 21.2%; Pred. No. 56;  
Matches 28; Conservative 24; Mismatches 52; Indels 28; Gaps 5;  
QY 197 WKLIYMP----RMCANSSSLKIPVYLAELHSAQGLNFVHFAKSSFYDDIFTGWIAOKL 252  
DB 3250 WSLMYFHRDRDLAANAICSAVPSHWVTSRTWS---IHASHEWMTTDMLTVM---N 3302

QY 253 KTHLLAQTWKKQKQLPSNCSLPYHVNIKSIQVTSKSYFSSRODHKSKVCVSIKGSANRW 312  
DB 3303 KMWILENPWMDKTPVESWEEIP-----YLGKRED--QWCGSLIGLTSRA 3345  
QY 313 TCIGDLNRSLSHQ 324  
DB 3346 TWAKNIQTAINQ 3357

Search completed: October 14, 2004, 00:29:05  
Job time : 26.3966 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 23:51:40 ; Search time 56.2616 Seconds  
(without alignments)  
1985.252 Million cell updates/sec

Title: US-10-790-589-2

Perfect score: 1927

Sequence: 1 MTAKPLRTVLSLFFALSGV.....KNHYIQAFHKLRLRYGPKC 354

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1927	100.0	354	11 Q9QY48	Q9QY48 mus musculus
2	1924	99.8	354	11 Q8C589	Q8C589 mus musculus
3	1373	71.3	356	11 Q9QZK9	Q9QZK9 rattus norv
4	1293.5	67.1	357	4 Q9NQW3	Q9NQW3 homo sapien
5	1293.5	67.1	361	4 Q8WZ79	Q8WZ79 homo sapien
6	692.5	35.9	375	12 Q9Q759	Q9Q759 fowlpox vir
7	567	29.4	153	4 Q8WZ80	Q8WZ80 homo sapien
8	566	29.4	350	11 Q9QZK8	Q9QZK8 rattus norv
9	455	23.6	232	12 Q9QZK8	Q9QZK8 rattus norv
10	444	23.0	366	5 Q9VED8	Q9VED8 drosophila
11	169.5	8.8	338	5 Q9B1L8	Q9B1L8 trichinella
12	151.5	7.9	339	5 Q8GQP5	Q8GQP5 trichinella
13	148.5	7.7	342	5 Q8GQP4	Q8GQP4 trichinella
14	111.5	5.8	832	10 Q7XQ00	Q7XQ00 oryza sativ
15	106	5.5	622	2 Q45155	Q45155 bacteroides
16	106	5.5	1115	12 Q9QMH2	Q9QMH2 parvo-like

17	104	5.4	276	5	Q962V2	trichinella
18	104	5.4	344	5	Q27073	trichinella
19	101.5	5.3	315	5	Q27076	trichinella
20	99	5.1	735	16	Q8XNF9	clostridium
21	97.5	5.1	1886	2	Q7X504	leptospora
22	97.5	5.1	2499	2	Q8GFA0	photobabdu
23	97	5.0	344	5	Q8MTE3	trichinella
24	97	5.0	419	10	Q9SMI9	trichinella
25	97	5.0	646	5	Q8ICY6	cuphea lanc
26	95	4.9	508	16	Q81ZA6	plasmodium
27	95	4.9	1770	5	Q813A3	bacillus an
28	94	4.9	172	2	Q9L910	plasmodium
29	94	4.9	2165	12	O99721	escherichia
30	94	4.9	2165	12	O41355	respiratory
31	94	4.9	2165	12	Q91W88	human respi
32	94	4.9	2165	12	Q82027	human respi
33	94	4.9	2165	12	O41356	human respi
34	94	4.9	2165	12	O82021	human respi
35	94	4.9	2165	12	P90197	human respi
36	93.5	4.9	712	12	Q9DUC1	human respi
37	93.5	4.9	716	17	Q9HLF9	tt virus. o
38	93	4.8	756	16	Q8DKB1	thermoplasm
39	92.5	4.8	320	12	Q9Q14	synchococc
40	91.5	4.7	399	17	Q8ZVF1	shope fibro
41	91.5	4.7	774	16	Q92CV7	pyrobaculum
42	91	4.7	763	12	Q9A78	listeria in
43	90.5	4.7	406	10	Q94147	tt virus. o
44	90.5	4.7	518	2	O87771	oryza sativ
45	90	4.7	504	2	Q9ZTP5	legionella
						rhizobium 1

## ALIGNMENTS

### RESULT 1

Q9QY48	ID	Q9QY48	PRELIMINARY;	PRT;	354 AA.
AC	Q9QY48;				
DT	01-MAY-2000	(TREMBLrel. 13, Created)			
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)			
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)			
DE	Deoxyribonuclease DLAD.				
GN	DNASE2B OR DLAD.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=C57BL/6; TISSUE=Liver;				
RX	MEDLINE=99428663; PubMed=10497274;				
RA	Shiokawa D., Tanuma S.;				
RT	"DLAD, a novel mammalian divalent cation-independent endonuclease with				
RL	homology to DNase II.";				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=C57BL/6;				
RX	MEDLINE=21556924; PubMed=11700027;				
RA	Shiokawa D., Tanuma S.;				
RT	"Isolation and Characterization of the DLAD/Dlad Genes, Which Lie				
RL	Head-to-Head with the Genes for Urate Oxidase.";				
DR	Biochem. Biophys. Res. Commun. 288:1119-1128(2001).				
DR	EMBL; AF128888; AAF05082.1; -				
DR	EMBL; AF334608; AAL34450.1; -				
DR	EMBL; AF334603; AAL34450.1; JOINED.				
DR	EMBL; AF334604; AAL34450.1; JOINED.				
DR	EMBL; AF334605; AAL34450.1; JOINED.				
DR	EMBL; AF334606; AAL34450.1; JOINED.				
DR	EMBL; AF334607; AAL34450.1; JOINED.				
DR	MGD; MGI:1913283; Dnaase2b.				
DR	GO; GO:0005737; C:cytoplasm; IDA.				
DR	GO; GO:0005576; C:extracellular; IDA.				



```

Db 301 SKWCYSTKDSQARWTCIGDLNRSHPQALRSRGFGICSKNRYIYQSFDRLVSHYASC 355
|||||
RESULT 4
Q9NQW3 PRELIMINARY; PRT; 357 AA.
AC Q9NQW3
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Deoxyribonuclease II beta.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=21272509; PubMed=11376952;
RA Krieser R.J., Maclellan K.S., Park J.P., Eastman A.;
RT "The cloning, genomic structure, localization, and expression of human
RL deoxyribonuclease II beta.";
RL Gene 269:205-216(2001).
DR EMBL; AF274571; AF276893.1; -.
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR004947; DNase_II.
DR Pfam; PF03265; DNase_II; 1.
SQ SEQUENCE 357 AA; 41197 MW; 6565D2036BD20955 CRC64;

Query Match 67.1%; Score 1293.5; DB 4; Length 357;
Best Local Similarity 66.4%; Pred. No. 2.1e-109;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

QY 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFYFKLPKRTSKASBEAGLQYL 60
|||
Db 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFYFKLPKRTSKASBEAGLQYL 60
|||
QY 61 YLDSTRQTNWKSLLYLNSTRSALGRITLQHLDTNHS--TNDTAYLIYNDGVPGSVNYSROY 119
|||||
Db 61 YLDSTRQTNWKSLLYLNSTRSALGRITLQHLDTNHS--TNDTAYLIYNDGVPGSVNYSROY 119
|||||
QY 120 GHAKGLLVNRTQGFWLHSPKPPV--HGVEYPTSGRRYGTGICITFGYSQFBEIDF 177
|||||
Db 120 GHAKGLLVNRTQGFWLHSPKPPV--HGVEYPTSGRRYGTGICITFGYSQFBEIDF 177
|||||
QY 121 GHTKGLLVNRTQGFWLHSPKPPV--HGVEYPTSGRRYGTGICITFGYSQFBEIDF 180
|||||
Db 121 GHTKGLLVNRTQGFWLHSPKPPV--HGVEYPTSGRRYGTGICITFGYSQFBEIDF 180
|||||
QY 178 QLLVLPNIYSCFIPSTFHWKLIYMPKANSLLKIPVRYLAELHSAQGLNPFVPAKSS 237
|||||
Db 178 QLLVLPNIYSCFIPSTFHWKLIYMPKANSLLKIPVRYLAELHSAQGLNPFVPAKSS 237
|||||
QY 181 QLLVLPNIYSCFIPSTFHWKLIYMPKANSLLKIPVRYLAELHSAQGLNPFVPAKSS 240
|||||
Db 181 QLLVLPNIYSCFIPSTFHWKLIYMPKANSLLKIPVRYLAELHSAQGLNPFVPAKSS 240
|||||
QY 238 FYTDDIFTGWIAQKLTLLAQTWQKKOELPNSCLPYHVYNIKSIYGVTSKYFSSROD 297
|||||
Db 238 FYTDDIFTGWIAQKLTLLAQTWQKKOELPNSCLPYHVYNIKSIYGVTSKYFSSROD 297
|||||
QY 241 SFLLDDIFAAMWAQKLTLLAQTWQKKOELPNSCLPYHVYNIKSIYGVTSKYFSSROD 300
|||||
Db 241 SFLLDDIFAAMWAQKLTLLAQTWQKKOELPNSCLPYHVYNIKSIYGVTSKYFSSROD 300
|||||
QY 298 HSKWCVSIKGSANRWTCIGDLNRSHPQALRSRGFGICSKNRYIYQSFDRLVSHYASC 354
|||||
Db 298 HSKWCVSIKGSANRWTCIGDLNRSHPQALRSRGFGICSKNRYIYQSFDRLVSHYASC 354
|||||
QY 301 HAKWCISQKGTNRWTCIGDLNRSHPQALRSRGFGICSKNRYIYQSFDRLVSHYASC 357
|||||
Db 301 HAKWCISQKGTNRWTCIGDLNRSHPQALRSRGFGICSKNRYIYQSFDRLVSHYASC 357
|||||

RESULT 5
Q9WZ79 PRELIMINARY; PRT; 361 AA.
AC Q9WZ79
DT 01-NAR-2002 (TrEMBLrel. 20, Created)
DT 01-NAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Endonuclease DLAD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=21556924; PubMed=11700027;
RA Shikawa D., Tanuma S.;
RT "Isolation and Characterization of the DLAD/DIAD Genes, Which Lie
RT Head-to-Head with the Genes for Urate Oxidase.";
RL Biochem. Biophys. Res. Commun. 288:1119-1128(2001).
DR EMBL; AF334602; AAL34449.1; -.
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR004947; DNase_II.
DR Pfam; PF03265; DNase_II; 1.
SQ SEQUENCE 361 AA; 41712 MW; C6FDD3F58F62CAC0 CRC64;

Query Match 67.1%; Score 1293.5; DB 4; Length 361;
Best Local Similarity 66.4%; Pred. No. 2.1e-109;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

QY 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFYFKLPKRTSKASBEAGLQYL 60
|||
Db 5 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFYFKLPKRTSKASBEAGLQYL 64
|||||
QY 61 YLDSTRQTNWKSLLYLNSTRSALGRITLQHLDTNHS--TNDTAYLIYNDGVPGSVNYSROY 119
|||||
Db 65 YLDSTRQTNWKSLLYLNSTRSALGRITLQHLDTNHS--TNDTAYLIYNDGVPGSVNYSROY 124
|||||
QY 120 GHAKGLLVNRTQGFWLHSPKPPV--HGVEYPTSGRRYGTGICITFGYSQFBEIDF 177
|||||
Db 125 GHTKGLLVNRTQGFWLHSPKPPV--HGVEYPTSGRRYGTGICITFGYSQFBEIDF 184
|||||
QY 178 QLLVLPNIYSCFIPSTFHWKLIYMPKANSLLKIPVRYLAELHSAQGLNPFVPAKSS 237
|||||
Db 185 QLLVLPNIYSCFIPSTFHWKLIYMPKANSLLKIPVRYLAELHSAQGLNPFVPAKSS 244
|||||
QY 238 FYTDDIFTGWIAQKLTLLAQTWQKKOELPNSCLPYHVYNIKSIYGVTSKYFSSROD 297
|||||
Db 245 SFLLDDIFAAMWAQKLTLLAQTWQKKOELPNSCLPYHVYNIKSIYGVTSKYFSSROD 304
|||||
QY 298 HSKWCVSIKGSANRWTCIGDLNRSHPQALRSRGFGICSKNRYIYQSFDRLVSHYASC 354
|||||
Db 305 HAKWCISQKGTNRWTCIGDLNRSHPQALRSRGFGICSKNRYIYQSFDRLVSHYASC 361
|||||

RESULT 6
O90759 PRELIMINARY; PRT; 375 AA.
AC O90759
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Deoxyribonuclease II (EC 3.1.22.1).
GN CEL1/DNASEII.
OS Fowlpox virus (FPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OX NCBI_TaxID=10261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FP-9;
RX MEDLINE=98325194; PubMed=9658122;
RA Laidlaw S.M., Anwar M.A., Thomas W., Green P., Shaw K., Skinner M.A.;
RT "Fowlpox virus encodes non-essential homologs of cellular alpha-SNAP,
RT PC-1 and an orphan human homolog of a secreted nematode protein.";
RL J. Virol. 72:6742-6751(1998).
DR EMBL; AJ006408; CAA07012.1; -.
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR004947; DNase_II.
DR Pfam; PF03265; DNase_II; 1.
KW Hydrolase.
SQ SEQUENCE 375 AA; 43064 MW; 36C641F74BAD25F6 CRC64;

Query Match 35.9%; Score 692.5; DB 12; Length 375;

```

Best Local Similarity 41.8%; Pred. No. 1.3e-54;		PRT; 350 AA.	
Matches 141; Conservative 57; Mismatches 120; Indels 19; Gaps 9;			
QY	25 EISCRNRYGAVDWFYFKLPKRTSKASEBAGLYLDSTRTQPNKSLYLINSTRSALG 84		
DB	46 KVCNVEGEVDWYFYFKLPKQKLGTK--GNEYLYIDSNPNKWKRGKVPINSRYSIIG 103		
QY	85 RTLOHLYDNTNSTNDTAYLYINDGVPQSVNRYOYGHAKGLLVN--RTQGFWLHHSVPK 142		
DB	104 KTLPIYDLVDS-KYIEYIFNDGIPGSKNYSSKVGHTKGYMAWNSDSVTGFWLHHSVPR 162		
QY	143 FP--PVHGYEYPTSGRRYGOTGICITCY-SQPEIDFQLLVLPNIYSCFIPFTFWKL 199		
DB	163 FPPSPVLGYNPYSGYVYQSMCLINLDYKGLGTALONTLPVNNPNVNC---SVTNKQL 219		
QY	200 IYPRMNCANSSSLKIPVRYLAELHSAQGLNFVHFAKSFYDDIDFTGWIAQKLKTHLLAQ 259		
DB	220 NNLYHLCNDKNYTYLYKNVSRMESRGEKELTFKASKYFRHDIMSAGWIGTLESLLSE 279		
QY	260 TWQKKQELPNCSLPHVHYNIKSGVTSKYFSRQDHSKWCYSI---KGSANRWTCIG 316		
DB	280 TWQRGSMITNCSSKHVHNKISINNVGTS-FINYDHSKWIYSLYDKRG---WVCIG 334		
QY	317 DLNBSLHQALRGGFICTKNHYIYQAFHKLRYLRYGFC 353		
DB	335 DINRSPQRHRRGGYACTRNGYLFKLKKTETVIEYEGC 371		
RESULT 7		PRT; 153 AA.	
ID	Q8WZ80 PRELIMINARY;		
AC	Q8WZ80; 01-MAR-2002 (TReMBLrel. 20, Created)		
DT	01-MAR-2002 (TReMBLrel. 20, Last sequence update)		
DT	01-JUN-2003 (TReMBLrel. 24, Last annotation update)		
DE	Lung-specific DLAD.		
OS	Homo sapiens (Human)		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21556924; PubMed=11700027;		
RA	Shlokawa D., Tanuma Si S.;		
RT	"Isolation and Characterization of the DLAD/Dlad Genes, Which Lie		
RT	Head-to-Head with the Genes for Urate Oxidase.";		
RL	Biochem. Biophys. Res. Commun. 288:1119-1128(2001).		
DR	EMBL; AF333389; AAL34448.1;		
DR	GO; GO:0004531; F:deoxyribonuclease II activity; IEA.		
DR	GO; GO:0006259; P:DNA metabolism; IEA.		
DR	InterPro; IPR004947; DNase_II.		
DR	Pfam; PF03265; DNase_II; 1.		
SQ	SEQUENCE 153 AA; 17816 MW; ABF4ED379AFA9FCF CRC64;		
Query Match		29.4%; Score 567; DB 4; Length 153;	
Best Local Similarity 66.7%; Pred. No. 1.1e-43;			
Matches 102; Conservative 18; Mismatches 33; Indels 0; Gaps 0;			
QY	202 MPRMNCANSSSLKIPVRYLAELHSAQGLNFVHFAKSFYDDIDFTGWIAQKLKTHLLAQTW 261		
DB	1 MPOLTRASSSEIPGRLLTTLQSAQKFLHFAKSDSFLDDIFAAWMAQRLKTHLLTETW 60		
QY	262 QKKQELPNCSLPHVHYNIKSGVTSKYFSRQDHSKWCYSIYKGSANRWTCIGDLNKS 321		
DB	61 QKRQELPNCSLPHVHYNIKAIRLSRHSYFSYQDHAKWCISQKGTNRVTCIGDLNKS 120		
QY	322 LHOALRGGFICTKNHYIYQAFHKLRYLRYGFC 354		
DB	121 PHQAFRSGGFICTQNWQIYQAFGLVLYYESCK 153		
RESULT 8			
Q9QZK8			
Best Local Similarity 29.4%; Score 566; DB 11; Length 350;			
Best Local Similarity 36.9%; Pred. No. 3.9e-43;			
Matches 127; Conservative 55; Mismatches 148; Indels 14; Gaps 8;			
QY	11 SLLFPALSGVLGTPEISCRNEYGEADVWFYFKLPKRTSKASEE-AGIQLVLDSTRTW 69		
DB	6 SLLAALLWV-PARALSCYGDGSRPVDWVYVYKLPANSGGDKPWKGLMYKMDQNSGW 64		
QY	70 NKSLYLINSTRSALGRTLQHLIYDTHNSTNDTAYLYINDGVP--GSVNYSRQYGHAKGLLV 127		
DB	65 QDGVHIDSKDCAVGLTLQPLQYNSS--QLAFLYNDQPKSSAQDSSSRGHGKTVLL 122		
QY	128 WNRTQGFWLHHSVPKFP-PVH--GYEYTSRRYGTGTCITGYSQFEEDFQLLVLP 184		
DB	123 LDQEGGFWLHHSVPKFPSPASSGAYSWPNARTYGTLLCVSLPSPFGIGKQLTYTP 182		
QY	185 NIYSCFIPSTFWKLIYMPRCANSSSLKIPVRYLAELHSAQGLNFVHFAKSFYDDIDF 244		
DB	183 LVYDHKLEGIFAOKLPDLSEVTKGHVLRERPNWSSVILTSRAGTTFQSFAPKFKGDDLY 242		
QY	245 TGWTAOKLTKTHLLAQTWQKQKQELPNSCLSPYHVNYIKSIGVTSKS--YFSRQDHSKWC 302		
DB	243 SGNLAAALGTNLQVQFWNSGILPNSCGTHKILDVTETGTFPGSPGFENATEDHSKWC 302		
QY	303 VSIKGSANRWTCIGDLNRSIHOALRGGFICTKNHYIYQAFHKL 346		
DB	303 VAPEGP---WVCVGMNKNKRETHRGGTLCTQVPAWKAFLSL 343		
RESULT 9			
Q9J5H0			
ID	Q9J5H0 PRELIMINARY;	PRT; 232 AA.	
AC	Q9J5H0; 01-OCT-2000 (TReMBLrel. 15, Created)		
DT	01-OCT-2000 (TReMBLrel. 15, Last sequence update)		
DT	01-JUN-2003 (TReMBLrel. 24, Last annotation update)		
DE	ORF FPV032 DNase II.		
GN	FPV032.		
OS	Fowlpox virus (FVP).		
OC	Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;		
OC	Avipoxvirus.		
OX	NCBI_TaxID=10261;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=20193820; PubMed=10729156;		
RA	Afonso C.L., Tulman B.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;		
RT	"The genome of fowlpox virus.";		
RL	J. Virol. 74:3815-3831(2000).		
RN	[2]		

Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., M.G., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirskas R., Tector C., Turner E., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
"The genome sequence of *Drosophila melanogaster*.";  
Science 287:2185-2195(2000).  
[2]  
SEQUENCE FROM N.A.  
STRAIN=Berkeley;  
Scapleton M., Brokstein P., Hong L., Agbayani A., Carlson J., Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E., George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G., Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S., Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M., Celinker S.;  
Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
EMBL; AE003720; AAF55486.1; -  
EMBL; AY075328; AAL68195.1; -  
FlyBase; FBgn0000477; DNaseII.  
GO; GO:0004531; Fdeoxyribonuclease II activity; NAS.  
DR InterPro; IPR004947; DNase II.  
DR Pfam; PF03265; DNase II; 1.  
DR SQU SEQUENCE 366 AA; 41340 MW; 8508BC7C7E86276A CRC64;  
Query Match 23.0%; Score 444; DB 5; Length 366;  
Best Local Similarity 31.0%; Pred. NO. 5.3e-32;  
Matches 107; Conservative 53; Mismatches 165; Indels 20; Gaps 7;  
QY 25 EISCRNEYGEAVDFWIFYKLPK---RTSKASEAGQVLYLDSTR-QTNKSLYLINSTR 80  
Db 22 KVSCKDEAGNDVDMWHLKPKHYQHNDLKGDTSLGLKLYVTSQNYDTWQMSKGFISDPL 81  
QY 81 SALGRTLOHLYTHNSTNTAVLIYNDVPGSVNYSRQYGHAKGLVWNRTOQFWLIHSV 140  
Db 82 SLPAQTLNPLND---DPSHTLLAAVNDQPPNGTVFS-SGSHAKGVVASDGETAIVIVHSV 137  
QY 141 PKFPPVHGVEYPTSGRRYGTGTCITFGYSQSEEDFQLLVLPQNYISCFISTFWKLI 200  
Db 138 PKFPTIDYSYPTSGEQVQAQSMCLVTLKGDEKQGLVLYNEPHFYORNPPLATSDSL 197  
QY 201 Y-MPRMCANSSSLKIPVRYLAELHSAQGLNFVHFSAKSYFYDDITPTGIAQKLTLLA 258  
Db 198 FPSLERALHGQWRTSEFPQKDLVRSLDGKKFLFGKSGRANVELYADVVPATLDVSLFV 257  
QY 259 QTWQKKQKQLPNCSLPHVIVNKKIGVTSKSY-FSSRODHKWCVS-----IKGS 308  
Db 258 EAWRDGAGNLPSCDKSKVLNVEGISNPELSVDFPKTQDHSKQWVRPTGILYVHVRVG 317  
QY 309 ANRWTCIGDNLNLSHQALRGGGFICTKNHYIYQAFHKLRYGFC 353  
Db 318 GGDWICVGDINRQEQQLHRRGGTVCHKARSVENLRYQLVTNYDKC 362  
RESULT 11  
Q9BJL8 PRELIMINARY; PRT; 338 AA.  
ID Q9BJL8  
AC Q9BJL8  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
DE Newborn larvae-specific protein S81.  
OS Trichinella spiralis (Trichina worm).  
OC Eukaryota; Metazoa; Nematoda; Enoplea; Trichocephalida;  
OC Trichinellidae; Trichinella.  
NCBI TaxID=6334;  
SEQUENCE FROM N.A.  
Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;  
Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
EMBL; AF198100; AAF44376.1; -  
DR GO; GO:0004531; Fdeoxyribonuclease II activity; IEA.  
DR GO; GO:0006259; P:DNA metabolism; IEA.  
DR InterPro; IPR004947; DNase II.  
DR Pfam; PF03265; DNase II; 1.  
DR SQU SEQUENCE 232 AA; 26772 MW; 14C05BB3BA79C73A CRC64;  
Query Match 23.6%; Score 455; DB 12; Length 232;  
Best Local Similarity 40.7%; Pred. NO. 2.9e-33;  
Matches 96; Conservative 39; Mismatches 85; Indels 16; Gaps 7;  
QY 126 LVVN-RTQGFWLHSPKFP--PVHGXYPTSGRRYGTGTCITFGY-SQFEIDFQLL 180  
Db 1 MAMNSDVTGFWLHSPKFP--PVHGXYPTSGRRYGTGTCITFGY-SQFEIDFQLL 60  
QY 181 VLPQNIYSCPTSPFTHFKLIWMPRCANSSSLKIPVRYLAELHSAQGLNFVHFSAKSYFT 240  
Db 61 VNNPNVYNC---SVTNKLNANLHLCNDKNYTLIKNVSRRKGEKFLTFKSKYFR 117  
QY 241 DDIFGTWIAQKLKTHLAQWQKKQBLPNSCLPXYHVNKISGVTSKVSFESSRQDHSK 300  
Db 118 HDIMSAMIGTLESLLSEWQRGESWITNCSKYHVNKISGVTSKVSFESSRQDHSK 176  
QY 301 WCYSI---KGSANRWTCIGDNLNLSHQALRGGGFICTKNHYIYQAFHKLRYGFC 353  
Db 177 WIVSLYDKKG---WVCIGDINRSPTRHGGVYACTRNGYLPKLTETVIEYEGC 228  
RESULT 10  
Q9VED8 PRELIMINARY; PRT; 366 AA.  
ID Q9VED8  
AC Q9VED8  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
DE C57780 protein (GH10876p).  
DR DNASEII OR C57780.  
OS *Drosophila melanogaster* (Fruit fly).  
NC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
NCBI\_TaxID=7227;  
SEQUENCE FROM N.A.  
STRAIN=Berkeley;  
MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Anantides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Franckoch C., Baldwin D.,  
RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Beran B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,







***This Page Blank (uspto)***

[illegible]

181 AAAAGACAAACAAAGGAAGTGGAGAGACTGGGTAGACTACCTGTACCTAGACTCTACA 240  
 181 AAAAGACAAACAAAGGAAGTGGAGAGACTGGGTAGACTACCTGTACCTAGACTCTACA 240  
 241 ACTAGAACTGGAGAGAGTGGAGAGTAACTAATGATGACACCAAGAGTGTGTTGGGAAGG 300  
 241 ACTAGAACTGGAGAGAGTGGAGAGTAACTAATGATGACACCAAGAGTGTGTTGGGAAGG 300  
 301 ACATTACACAGCTATATGAGCATATGCCCTTAAGAGTAAACACACAGCTATCTAATA 360  
 301 ACATTACACAGCTATATGAGCATATGCCCTTAAGAGTAAACACACAGCTATCTAATA 360  
 361 TACATGATGAGTCCCTTAACCTGTGAATACAGTAAAGTATGACACCAAGAGT 420  
 361 TACATGATGAGTCCCTTAACCTGTGAATACAGTAAAGTATGACACCAAGAGT 420  
 421 TTACTGTGTGGAACAGAGTTCAGAGGTTCTGGCTGATTCATTCATCCCTCAGTTTCCT 480  
 421 TTACTGTGTGGAACAGAGTTCAGAGGTTCTGGCTGATTCATTCATCCCTCAGTTTCCT 480  
 481 CCAATTCGGGAAGAGCTATGATATCCACCCACAGGAGACGAATGGAACAAAGTGC 540  
 481 CCAATTCGGGAAGAGCTATGATATCCACCCACAGGAGACGAATGGAACAAAGTGC 540  
 541 ATCTGCATACCTTTCAAGTACAAACAGTATGAGCAATAGATTCTCAGCTCTTGTCTGC 600  
 541 ATCTGCATACCTTTCAAGTACAAACAGTATGAGCAATAGATTCTCAGCTCTTGTCTGC 600  
 601 AACCCCAAGCTATAGTCTCCATCCAGCCACCTTTTACCAGGAGCTCATTCACATG 660  
 601 AACCCCAAGCTATAGTCTCCATCCAGCCACCTTTTACCAGGAGCTCATTCACATG 660  
 661 CCCAGCTGTGACACAGGAGCTATGATATCCACCCACAGGAGTTCCTGAGGCTCTCACCACAT 720  
 661 CCCAGCTGTGACACAGGAGCTATGATATCCACCCACAGGAGTTCCTGAGGCTCTCACCACAT 720  
 721 CAGTCGGCCAGGACAAATTCCTCCATTTGCAAGTGGATTCCTTTCTTGCAGAC 780  
 721 CAGTCGGCCAGGACAAATTCCTCCATTTGCAAGTGGATTCCTTTCTTGCAGAC 780  
 781 ATCTTTGAGCTGTGATGGCTCAACGGCTGAAGACACACTTTGTTAAACAGAACTGGCAG 840  
 781 ATCTTTGAGCTGTGATGGCTCAACGGCTGAAGACACACTTTGTTAAACAGAACTGGCAG 840  
 841 CGAAAAACAGAGAGCTTCCTTCAAACTGCTCCCTTCTTACCATGCTCAATATATAA 900  
 841 CGAAAAACAGAGAGCTTCCTTCAAACTGCTCCCTTCTTACCATGCTCAATATATAA 900  
 901 GCATTAATATATACGACACTCTTATTTTCACTTCTTATCAAGATCAAGCAAGTGGT 960  
 901 GCATTAATATATACGACACTCTTATTTTCACTTCTTATCAAGATCAAGCAAGTGGT 960  
 961 ATTTCCCAAGGACCAAAATCGCTGGACATGTTGGAGACTTAATCGGAGTCA 1020  
 961 ATTTCCCAAGGACCAAAATCGCTGGACATGTTGGAGACTTAATCGGAGTCA 1020  
 1021 CACCAAGCTTCAGAGTGGAGTTCATTTGATCCAGAAATGCAATTTACCAAGCA 1080  
 1021 CACCAAGCTTCAGAGTGGAGTTCATTTGATCCAGAAATGCAATTTACCAAGCA 1080  
 1081 TTTCAAGATAGTATATATATATGAAAGCTGTAAAGTAACTTTGGTGAAGGACACAGT 1140  
 1081 TTTCAAGATAGTATATATATGAAAGCTGTAAAGTAACTTTGGTGAAGGACACAGT 1140  
 1141 ACTATCAATGAAACCTTGACAAATGGTCTTCTTCAATPACACCTTTCTTATTTTAA 1200  
 1141 ACTATCAATGAAACCTTGACAAATGGTCTTCTTCAATPACACCTTTCTTATTTTAA 1200  
 1201 GGCCTGTGAATATATATATATGAAAGCTGTAAAGTAACTTTGGTGAAGGACACAGT 1260  
 1201 GGCCTGTGAATATATATATGAAAGCTGTAAAGTAACTTTGGTGAAGGACACAGT 1260  
 1261 ACCATTATA 1268

Db 1261 ACCATTATA 1268  
 RESULT 2  
 US-10-240-709-3  
 ; Sequence 3, Application US/10240709  
 ; Publication No. US20030212023A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Eastman, Alan R.  
 ; APPLICANT: Trustees of Dartmouth College  
 ; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND cDNAs  
 ; FILE REFERENCE: DC-0154  
 ; CURRENT APPLICATION NUMBER: US/10/240,709  
 ; CURRENT FILING DATE: 2002-10-01  
 ; PRIOR APPLICATION NUMBER: 09/541,840  
 ; PRIOR FILING DATE: 2000-04-03  
 ; PRIOR APPLICATION NUMBER: 09/574,942  
 ; PRIOR FILING DATE: 2000-05-19  
 ; NUMBER OF SEQ ID NOS: 6  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 3  
 ; LENGTH: 1268  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-10-240-709-3  
 Query Match 100.0%; Score 1268; DB 13; Length 1268;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1268; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGGGGAAAGTCTCTGCTGGCATGAAATAAATGAAACAGAAATGATGCGAAGACTG 60  
 Db 1 ATGGGGAAAGTCTCTGCTGGCATGAAATAAATGAAACAGAAATGATGCGAAGACTG 60  
 QY 61 CTAAAGACATCTTTGCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 120  
 Db 61 CTAAAGACATCTTTGCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 120  
 QY 121 ATTTCTGCAAAATGAAGAGGAAAGTGGAGAGTGGGTAGAGTACCTGTACCTAGACTCTACA 180  
 Db 121 ATTTCTGCAAAATGAAGAGGAAAGTGGAGAGTGGGTAGAGTACCTGTACCTAGACTCTACA 180  
 QY 181 AAAAGACAAACAAAGGAAGTGGAGAGTGGGTAGAGTACCTGTACCTAGACTCTACA 240  
 Db 181 AAAAGACAAACAAAGGAAGTGGAGAGTGGGTAGAGTACCTGTACCTAGACTCTACA 240  
 QY 241 ACTAGAACTGGAGAGAGTGGAGCAACTAATGATGACACCAAGAGTGTGTTGGGAAGG 300  
 Db 241 ACTAGAACTGGAGAGAGTGGAGCAACTAATGATGACACCAAGAGTGTGTTGGGAAGG 300  
 QY 301 ACATTACACAGCTATATGAGCATATGCCCTTAAGAGTAAACACACAGCTATCTAATA 360  
 Db 301 ACATTACACAGCTATATGAGCATATGCCCTTAAGAGTAAACACACAGCTATCTAATA 360  
 QY 361 TACAATGATGAGTCCCTTAACCTGTGAATACAGTAAAGTATGACACCAAGAGT 420  
 Db 361 TACAATGATGAGTCCCTTAACCTGTGAATACAGTAAAGTATGACACCAAGAGT 420  
 QY 421 TTACTGTGTGGAACAGAGTTCAGAGGTTCTGGCTGATTCATTCATCCCTCAGTTTCCT 480  
 Db 421 TTACTGTGTGGAACAGAGTTCAGAGGTTCTGGCTGATTCATTCATCCCTCAGTTTCCT 480  
 QY 481 CCAATTCGGGAAGAGCTATGATATCCACCCACAGGAGACGAATGGAACAAAGTGC 540  
 Db 481 CCAATTCGGGAAGAGCTATGATATCCACCCACAGGAGACGAATGGAACAAAGTGC 540  
 QY 541 ATCTGCATACCTTTCAAGTACAAACAGTATGAGCAATAGATTCTCAGCTCTTGTCTGC 600  
 Db 541 ATCTGCATACCTTTCAAGTACAAACAGTATGAGCAATAGATTCTCAGCTCTTGTCTGC 600  
 QY 601 AACCCCAAGCTATAGTCTCCATCCAGCCACCTTTTACCAGGAGCTCATTCACATG 660

```

Db 601 AACCCCAACGCTATAGCTGCTCCATCCAGCCACTTCCACGAGGAGCTATTACATG 660
QY CCCAGAGCTGTCCACAGGCGCAGCTCATCAGAGATTCCTGCGAGGCTCCTCAGCAGACTT 720
Db 661 CCCAGAGCTGTCCACAGGCGCAGCTCATCAGAGATTCCTGCGAGGCTCCTCAGCAGACTT 720
QY 721 CAGTCGGCCACAGGAGCAAAAATTCCTCCATTTTGCAGAGTCCGATTCCTTTTTCGAGC 780
Db 721 CAGTCGGCCACAGGAGCAAAAATTCCTCCATTTTGCAGAGTCCGATTCCTTTTTCGAGC 780
QY 781 ATCTTTGAGAGCTGAGTGCCTCAACGCTGAAGCACAATTTGTTTAAACAGAACTGGCAG 840
Db 781 ATCTTTGAGAGCTGAGTGCCTCAACGCTGAAGCACAATTTGTTTAAACAGAACTGGCAG 840
QY 841 CGAAAAAGACAAGAGCTTCCTTCAAACTGCTCCCTTCCATGATGCTTACCAATATAAAA 900
Db 841 CGAAAAAGACAAGAGCTTCCTTCAAACTGCTCCCTTCCATGATGCTTACCAATATAAAA 900
QY 901 GCAATTAATATATACGACACTCTTATTTTCAATGCTTATCAAGATCAGCCCAAGTGTGT 960
Db 901 GCAATTAATATATACGACACTCTTATTTTCAATGCTTATCAAGATCAGCCCAAGTGTGT 960
QY 961 ATTTCCCAAGAGGCAACAAAATTCGCTGAGATGCTGAGACCTTAAATCGAGTCCA 1020
Db 961 ATTTCCCAAGAGGCAACAAAATTCGCTGAGATGCTGAGACCTTAAATCGAGTCCA 1020
QY 1021 CACCAAGCTTTCAGAGTGGAGGATTCATTTTACCCAGATTCGCAATTCGCAATTTACCAAGCA 1080
Db 1021 CACCAAGCTTTCAGAGTGGAGGATTCATTTTACCCAGATTCGCAATTCGCAATTTACCAAGCA 1080
QY 1081 TTTCAAGGATTAATATATATATGAAAGCTGTAAGTAAATTTGGTGAAGGACACAGGT 1140
Db 1081 TTTCAAGGATTAATATATATATGAAAGCTGTAAGTAAATTTGGTGAAGGACACAGGT 1140
QY 1141 ACTATCAATTGAAACCTTGCAATGGGTCTTCTTCCATTAACACTTCTTTTATTTTAAA 1200
Db 1141 ACTATCAATTGAAACCTTGCAATGGGTCTTCTTCCATTAACACTTCTTTTATTTTAAA 1200
QY 1201 GGCTGTGAAATATATATATATACCTGCAATTCATATACAAATATAAATATTTCTCATGTT 1260
Db 1201 GGCTGTGAAATATATATATATACCTGCAATTCATATACAAATATAAATATTTCTCATGTT 1260
QY 1261 ACCAATTA 1268
Db 1261 ACCAATTA 1268

```

RESULT 3

```

US-10-790-589-3
; Sequence 3, Application US/10790589
; Publication No. US20040142376A1
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Krieser, Ronald J.
; APPLICANT: Trustees of Dartmouth College
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCE: DC-0154
; CURRENT APPLICATION NUMBER: US/10/790,589
; CURRENT FILING DATE: 2004-03-01
; PRIOR APPLICATION NUMBER: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1268
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-790-589-3

```

Query Match

100.0%; Score 1268; DB 17; Length 1268;

```

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1268; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGGAAGTGTCTCTGCTGTGGCATGAAATAAATAAGAAACAGAAAATGATGCGAAGACTG 60
Db 1 ATGGGGAAGTGTCTCTGCTGTGGCATGAAATAAATAAGAAACAGAAAATGATGCGAAGACTG 60
QY 61 CTAAGAACATCTTTTGTCTTTGTCTTCTTCTTGGGCTCTTTGGGGTGTGGGGGCGAGCAACA 120
Db 61 CTAAGAACATCTTTTGTCTTTGTCTTCTTCTTGGGCTCTTTGGGGTGTGGGGGCGAGCAACA 120
QY 121 ATTTTCATGCAAGAAATGAAGAGGCAAGAGCTGTGCACTGCTTTACTTTTATAAGTTACCT 180
Db 121 ATTTTCATGCAAGAAATGAAGAGGCAAGAGCTGTGCACTGCTTTACTTTTATAAGTTACCT 180
QY 181 AAAAGACAAAACAAAGAAAGTGGAGAGACTGGGTGTAGAGTACCTGTACCTAGACTCTTACA 240
Db 181 AAAAGACAAAACAAAGAAAGTGGAGAGACTGGGTGTAGAGTACCTGTACCTAGACTCTTACA 240
QY 241 ACTAGAGCTGGAGGAAAGAGTGAAGCAACTAATGAATGACCAAGAGTGTTTTGGGAAGG 300
Db 241 ACTAGAGCTGGAGGAAAGAGTGAAGCAACTAATGAATGACCAAGAGTGTTTTGGGAAGG 300
QY 301 ACATTACAACAGCTATATATGAAGCATATGCTCTTAAGAGTAAACACACAGCCTATCTAATA 360
Db 301 ACATTACAACAGCTATATATGAAGCATATGCTCTTAAGAGTAAACACACAGCCTATCTAATA 360
QY 361 TACAATGATGGAGTCCCTTAAACCTGTGAATTAACAGTAGAAAGTATGACACACCAAGGT 420
Db 361 TACAATGATGGAGTCCCTTAAACCTGTGAATTAACAGTAGAAAGTATGACACACCAAGGT 420
QY 421 TTTACTGCTGTGAAACAGAGTTCAAGGGTCTGCTGCTGANTTCATTCCTCAGTTTCTT 480
Db 421 TTTACTGCTGTGAAACAGAGTTCAAGGGTCTGCTGCTGANTTCATTCCTCAGTTTCTT 480
QY 481 CCAATTCGGAAGAGCTATGATTAATCCACACAGGGAGAGCAAAATGACAAAGTGGC 540
Db 481 CCAATTCGGAAGAGCTATGATTAATCCACACAGGGAGAGCAAAATGACAAAGTGGC 540
QY 541 ATCTGATCAACTTTCAAGTAAACCAAGTATGAGGCAATAGATTCTCAGCTCTTTGGTCTGC 600
Db 541 ATCTGATCAACTTTCAAGTAAACCAAGTATGAGGCAATAGATTCTCAGCTCTTTGGTCTGC 600
QY 601 AACCCCAACGCTATAGCTGCTCCATCCAGCCACTTTCACAGAGTCAATTCACATG 660
Db 601 AACCCCAACGCTATAGCTGCTCCATCCAGCCACTTTCACAGAGTCAATTCACATG 660
QY 661 CCCAGCTGTGACAGGCGCAGCTCATCAGAGATTCCTGCGAGGCTCCTCACCACACTT 720
Db 661 CCCAGCTGTGACAGGCGCAGCTCATCAGAGATTCCTGCGAGGCTCCTCACCACACTT 720
QY 721 CAGTCGGCCACAGGAGCAAAAATTCCTCCATTTTGCAGAGTCCGATTCCTTTTTCGAGC 780
Db 721 CAGTCGGCCACAGGAGCAAAAATTCCTCCATTTTGCAGAGTCCGATTCCTTTTTCGAGC 780
QY 781 ATCTTTGAGAGCTGAGTGCCTCAACGCTGAAGCACAATTTGTTTAAACAGAACTGGCAG 840
Db 781 ATCTTTGAGAGCTGAGTGCCTCAACGCTGAAGCACAATTTGTTTAAACAGAACTGGCAG 840
QY 841 CGAAAAAGACAAGAGCTTCCTTCAAACTGCTCCCTTCCATGATGCTTACCAATATAAAA 900
Db 841 CGAAAAAGACAAGAGCTTCCTTCAAACTGCTCCCTTCCATGATGCTTACCAATATAAAA 900
QY 901 GCAATTAATATATACGACACTCTTATTTTCAATGCTTATCAAGATCAGCCCAAGTGTGT 960
Db 901 GCAATTAATATATACGACACTCTTATTTTCAATGCTTATCAAGATCAGCCCAAGTGTGT 960
QY 961 ATTTCCCAAGAGGCAACAAAATTCGCTGAGATGCTGAGACCTTAAATCGAGTCCA 1020
Db 961 ATTTCCCAAGAGGCAACAAAATTCGCTGAGATGCTGAGACCTTAAATCGAGTCCA 1020
QY 1021 CACCAAGCTTTCAGAGTGGAGGATTCATTTTACCCAGATTCGCAATTCGCAATTTACCAAGCA 1080

```



```
; NAME/KEY: CDS
; LOCATION: (213)..(1274)
US-10-670-863-2

Query Match      53.1%; Score 673.8; DB 17; Length 1652;
Best Local Similarity 75.1%; Pred. No. 1.1e-192;
Matches 884; Conservative 0; Mismatches 282; Indels 11; Gaps 3;

QY 37 AAACAGAAATGATGGCAAGACTGTAAAGACATCTTTGCTTTGCTCTCTCTTGGCCCTC 96
DB 204 AGATAGAAATGACAGCAAGCCCTTAAGACAGTTCTTTCTTTCTTTGCCCCCTC 263
QY 97 TTTGGGTGCTGGGGCAGCAACAAATTTATGCGAAGATGAGAGGAAAGCTGTGAC 156
DB 264 TCTGGGTGCTGGGGCAGCAACAAATCTCATGAGAAATGATGATGGTGAAGCTGTGAC 323
QY 157 TGGTTTACTTTTATAGTTTACCTTAAAGACAAACAAAGGAAAGTGGAGACTGGGTTA 216
DB 324 TGGTTTATCTTTTATAGTTTACCTTAAAGGAGTGAAGCAAGTGAAGGCGGGCTG 383
QY 217 GAGTACCTGTACCTAGACTCTACAACTAGAACTGGAGGAGAGTGAAGCACTAATGAAT 276
DB 384 CAGTACCTGTACCTAGACTCTACAACTAGAACTGGAGGAGAGTGAAGCACTAATGAAT 443
QY 277 GACACCAAGAGTGTGGGAGGAGCAATTAACAGCTATATGAGCATATGCTCTTAAG 336
DB 444 AGCACCAGGAGTGTCTGGGAGGAGCACTTACAGCATCTGTATGACACACATAATTCACG 503
QY 337 AGTAAACACACAGCTATCTAATATACATGAGTGGTCCCTTAACTTGAATTAAGT 396
DB 504 AATGACA--CAGCTATCTAATATACAGATGGTGTCTCTGATGTGGAATTAACG 560
QY 397 AGAAAGTATGACACACAAAGGTTTACTGTCTGTGAAACAGAGTTCAAGGTTCTGGCTG 456
DB 621 ATACACTCTGTCTCCAAAGTTTCCCAAGTTC-----ATGGCTATGAGTACCCAACTCG 674
QY 517 GCGAGAGAAATGACAAAGTGGCATCTGCATTAACATTTCAAGTACAAACAGTATGAGGCA 576
DB 675 GCGAGCGATATGGAACAAACCGGCATCTGCATCACTTTCCGATACAGCCAGTTTGAAGAA 734
QY 577 ATAGATTTTCAGCTCTTGTGCTGCAACCCCAAGTCTATAGCTGTCTCACTCCAGCCACC 636
DB 735 ATAGATTTTCAGCTCTTGTGCTTACAAACCAATCTAAGCTGTCTCACTCCAGCCACC 794
QY 637 TTTACACAGAGCTATTCATGCCCCAGCTGTGCAACGAGGCGCCAGCTCATCAGAGATT 696
DB 795 TTTCACTGGAACTTATCTACATGCCCCGAGTGTGTCACCACTTCCATTAAGATC 854
QY 697 CTGGCAGAGCTCTACACACACTTCACTGCGCCAGGAGCAAAATTCCTCAATTTGCA 756
DB 855 CTGTCCGTAACCTCTGTAACCTGCACTCAGCCCGAGGCTTAAACTTCGTCCATTTTGA 914
QY 757 AAGTCGGAATCTTTTCTTGACGACATCTTTTGACGCTGTGATGCTCAACGCTGAGACA 816
DB 915 AAATCAAGTTTTTATACTGATGACATCTTTA CAGATGATGATGATGATGATGATGATGAT 974
QY 817 CACTTTGTTAAGAACTCTGAGCGGAGGAAAGAAAGAGCTTCTTCAAACTGCTCCCTT 876
DB 975 CATTTGTTAGCAAAACCTCTGAGGAGGAAAGAAAGAGCTTCTTCAAACTGCTCCCTG 1034
QY 877 CTTTACCATGTCTCAATATAAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 936
DB 1035 CTTTACCATGTCTCAATATAAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1094
QY 937 TATCAAGATACGCGCAAGTGTGTATTTCCCAAAAGGCGCAACAAATTCGCTGGCAATG 996
DB 1095 CGCCAGACCAATTCAAATGTTGTGTTTCCATAAAGGGCTCGCAAAATCGCTGGACCTGC 1154
QY 997 ATTGGAGACCTAAATCGGAGTCCACACCAAGCTTTTCAAGTGGAGGATTCATTTTGTACC 1056

; Sequence 1, Application US/09949434
; Patent No. US20020028495A1
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND cDNAs
; FILE REFERENCE: DC-0137
; CURRENT APPLICATION NUMBER: US/09/949,434
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1224
; TYPE: DNA
; ORGANISM: Mus sp.
US-09-949-434-1

Query Match      53.1%; Score 672.8; DB 9; Length 1224;
Best Local Similarity 75.0%; Pred. No. 1.8e-192;
Matches 871; Conservative 0; Mismatches 282; Indels 9; Gaps 2;

QY 37 AAACAGAAATGATGGCAAGACTGTAAAGACATCTTTGCTTTGCTCTCTTGGCCCTC 96
DB 31 AGATAGAAATGACAGCAAGCCCTTAAGACAGTTCTTTCTTTGCTCTCTTGGCCCTC 90
QY 97 TTTGGGTGCTGGGGCAGCAACAAATTTATGCGAAGATGAGAGGAAAGCTGTGAC 156
DB 91 TCTGGGTGCTGGGGCAGCAACAAATCTCATGAGAAATGATGATGAGAGCTGTGAC 150
QY 157 TGGTTTACTTTTATAGTTTACCTTAAAGACAAACAAAGGAAAGTGGAGACTGGGTTA 216
DB 151 TGGTTTATCTTTTATAGTTTACCTTAAAGGAGTGAAGCAAGTGAAGGCGGGCTG 210
QY 217 GAGTACCTGTACCTAGACTCTACAACTAGAACTGGAGGAGAGTGAAGCACTAATGAAT 276
DB 211 GAGTACCTGTACCTAGACTCTACAACTAGAACTGGAGGAGGAGCTCTACCTGATTAAC 270
QY 277 GACACCAAGAGTGTGGGAGGAGCAATTAACAGCTATATGAGCATATGATGACACATATCTTAAG 336
DB 271 AGCACCAGGAGTGTCTGGGAGGAGCTTACAGCATCTGTATGACACATATCTTAAG 330
QY 337 AGTAAACACACAGCTATCTAATATACATGAGTGGTCCCTTAACTTGAATTAAGT 396
DB 331 AATG---ACAGAGCTATCTAATATACAAAGTGTCTCTCTGATCTGTGATTAACG 387
QY 397 AGAAAGTATGACACACAAAGGTTTACTGTGTGAGAGAGTTCAGGTTTCTGGCTG 456
DB 388 AGCAGATGAGCAATGCGCAAGGCTGTCTGTGATGAGGAGCAAGCAAGGCTTCTGGCTG 447
QY 457 ATTCAATCCATCCCTCAGTTTCTTCCAAATTCGGAAGAGGCTATGATTAATCCACCA 516
DB 448 ATACACTCTGTTCACCAAGTTTCCCAAGTTC-----ATGGCTATGAGTACCAACCTCG 501
```

517 GGAGAGCAAAATGCGAAGTGCATCTGCAATCTTCAAGTACACCAAGTATGAGCA 576  
 502 GGGAGGCGATATGACAAACCGGCGATCTGCATCACTTTCGGATACAGCCAGTTTGAGAA 561  
 577 ATAGATTTCTCAGCTCTTGGTCTGCAACCCCAACGCTCTATAGCTGCTCCATCCCGAGCCACC 636  
 562 ATAGATTTCTCAGCTCTTGGTCTTACACCAACCAATCTACAGCTCTTCAATCCCGAGCCACC 621  
 637 TTTCACGAGGAGCTCAATCTCATGCCCCAGCTGTCACCGAGGCGAGCTCATCATGAGATT 696  
 622 TTTCACGAGGAGCTCAATCTCATGCCCCAGCTGTCACCGAGGCGAGCTCATCATGAGATT 681  
 697 CTGCGAGGCTCTTCCACCACTTCACTGCGCCCGAGGAGCAAAATTCCTCCATTTTGCA 756  
 682 CTTGTCGGGTACCTTCTGAGTCACTGAGCCCGAGGCTTAACTCTTGCTCCATTTTGCA 741  
 757 AAGTCGGATTTCTTCTGAGGAGCATCTTTGCGAGCTGGATGGCTCAACGCGCTGAAGACA 816  
 742 AATCAAGTTTTTATATCTGATGACATCTTTACAGGATGGATGATCTCAAAAGTTGAAGACA 801  
 817 CACTTGTAAACAGAACTGCGAGCGGAAAGACAGAGCTTCTTCAAACTGCTCCCTT 876  
 802 CATTGTGTAGCAAACTGCGAGGAAAGACAGAGCTTCTTCAAACTGCTCCCTT 861  
 877 CTTTACCATGCTTACATATAAAGCAATTAATTTATCAGCACACTCTTATTTTCAAGTTCT 936  
 862 CTTTACCATGCTTACATATAAAGCAATTAATTTATCAGCACACTCTTATTTTCAAGTTCT 921  
 937 TATCAAGATCAGCGAAGTGTATTTCCCAAAAGGCGACCAAAATTCGCTGGACATGT 996  
 922 CGCAAGACCAATTCNAATGTGTGTCTTCCATAAAGGCTCCGCAATTCGCTGGACATGT 981  
 997 ATTGGAGACCTAAATCGGAGTCCACCAAGCTTCAAGAGTGTGAGGATTCATTTGTACC 1056  
 982 ATTGGAGACCTAAATCGGAGTCCACCAAGCTTCAAGAGTGTGAGGATTCATTTGTACC 1041  
 1057 CAGATTTGGCAAAATTTTACCAAGCATTTTCAAGGATTAGTATTATATCTATGAAAGCTGTAA 1116  
 1042 AAGAATCACTACATTTACCGGCAATTCATAAATTTATCTCCGTTATGGGTTCTGTAAG 1101  
 1117 TAACTTGGTGAAGGACACAGGTAATCTATCTTGAAGGCTTGAAGGCTTCTTCTTCC 1176  
 1102 TAACTTGGTGAAGGACACAGGTAATCTATCTTGAAGGCTTGAAGGCTTCTTCTTCC 1161  
 1177 ATTACACCTTCTTATATTTTA 1198  
 1162 CTTGGATCTGTTCTCCATAATA 1183

RESULT 7

US-10-240-709-1  
 ; Sequence 1, Application US/10240709  
 ; Publication No. US20030212023A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Eastman, Alan R.  
 ; APPLICANT: Krieser, Ronald J.  
 ; APPLICANT: Trustees of Dartmouth College  
 ; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS  
 ; FILE REFERENCE: DC-0154  
 ; CURRENT APPLICATION NUMBER: US/10/240,709  
 ; CURRENT FILING DATE: 2002-10-01  
 ; PRIOR APPLICATION NUMBER: 09/541,840  
 ; PRIOR FILING DATE: 2000-04-03  
 ; PRIOR APPLICATION NUMBER: 09/574,942  
 ; PRIOR FILING DATE: 2000-05-19  
 ; NUMBER OF SEQ ID NOS: 6  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO 1  
 ; LENGTH: 1224  
 ; TYPE: DNA  
 ; ORGANISM: Mus sp.  
 US-10-240-709-1

Query Match 53.1%; Score 672.8; DB 13; Length 1224;  
 Best Local Similarity 75.0%; Pred. No. 1.8e-192;  
 Matches 871; Conservative 0; Mismatches 282; Indels 9; Gaps 2;  
 37 AACAGAAATGATGGCAAGCTGCTTAAGACATCCCTTTGCTTGTCTTCTTGGCTC 96  
 31 AGATAGAAATGACAGCAAGCCCTCTAAGAACAGTTCTTCTTGTCTTCTTGGCTC 90  
 97 TTTGGGCTGCTGGGGCAGCAACAAATTTATGAGAAATGAAGAGGAAAGCTGTGGAC 156  
 91 TCTGGGCTGCTGGGGCAGCAACAAATTTATGAGAAATGAATATGTTGAAGCTGTGGAC 150  
 157 TGGTTTATCTTTTATATAGTTTACCTTAAAGACAAACAAAGAAAGTGAAGAGAGCTGTTA 216  
 151 TGGTTTATCTTTTATATAGTTTACCTTAAAGAGCTAGCAAGGCAAGTGAAGGCGGGCTG 210  
 217 GAGTACCTGTACCTAGTACTCTAGACTAGAGCTGGAGGAGAGTGAAGCACTAATGAAT 276  
 211 CAGTACCTGTACCTGGACTCCACAGACAAACCTTGGAAACAGAGCCTTACCTGATTAAC 270  
 277 GACACCAAGAGTGTTTTGGGAAGGACATTAACACAGCTATATGAAGCATATGCTCTTAAG 336  
 271 AGCACAGAGTGTCTTGGGAGGACCTTACAGCATCTGTATGACACACATAATTCACAG 330  
 337 AGTAAACACACAGCTTATCTAATATATACATGATGGAGTCCCTTAAACCTGTGAATACGT 396  
 331 AATG---ACACAGCTTATCTAATATACACAGTGTGTCTTGGATCTGTGAATTACAGC 387  
 397 AGAAAGTATGACACACCAAGGTTTACTGCTGTGGAAACAGAGTTCAGGGTCTTGGCTG 456  
 388 AGACAGTATGACATGCAAAAGGCTGCTGTGTATGAAACAGAACGAGGGGTTCTGGCTG 447  
 457 ATTCAATTCATCTCAGTTTCTTCCATTCCTGGAAGAGGCTATGATTTATCCACCCACA 516  
 448 ATACACTCTGTCTCCAGTTTCCCCAGTTC-----ATGCTATGAGTACCCACCTCG 501  
 517 GGGAGACAAATGACAAAGTGGCATCTGCATAATCTTCAAGTACCAACAGTATGAGCA 576  
 502 GGGAGGCGATATGACAAACCGGCACTCTCATCTTTCGGATACAGCCAGTTTGAAGAA 561  
 577 ATAGATTTCTCAGCTTGTGCTGCAACCCCAAGCTCTATAGCTGTCTTCCATCCAGCCACC 636  
 562 ATAGATTTCTCAGCTTGTGCTTACACCAACATCTACAGCTGTTTCTTCCAGAGCACC 621  
 637 TTTCACGAGGAGCTCAATTCATGCTCCAGCTGTGTCACAGGCGAGCTCATCAGAGATT 696  
 622 TTTCACGAGGAGCTCAATTCATGCTCCAGGAGTGTGTCACCTCCAGTTCTTAAAGATC 681  
 697 CTGCGAGGCTCTTCCACACATTCAGTGGGCGGAGGACAAATTCCTCCATTTGCA 756  
 682 CTTGTCGGTACCTCGCTGAACCTGCACTCAGCCAGGCTTAACTTCGTCTCATTTTGA 741  
 757 AAGTCGGATTTCTTCTGAGGAGCATCTTTCAGCTGCTGATGGCTCAACGCTGAAGACA 816  
 742 AAATCAAGTTTTTATCTGATGACATCTTTACAGGATGGATGATCTCAAAAGTTGAAGACA 801  
 817 CACTTGTAAACAGAACTGCGAGGAGGAAAGACAGAGCTTCTTCAAACTGCTCCCTT 876  
 802 CATTGTGTAGCAAACTGCGAGGAAAGACAGAGCTTCTTCAAACTGCTCCCTT 861  
 877 CTTTACCATGCTTACATATAAAGCAATTAATTTATCAGCACACTCTTATTTTCAAGTTCT 936  
 862 CTTTACCATGCTTACATATAAAGCAATTAATTTATCAGCACACTCTTATTTTCAAGTTCT 921  
 937 TATCAAGATCAGCGAAGTGTATTTCCCAAAAGGCGACCAAAATTCGCTGGACATGT 996  
 922 CGCAAGACCAATTCNAATGTGTGTCTTCCATAAAGGCTCCGCAATTCGCTGGACATGT 981  
 997 ATTGGAGACCTAAATCGGAGTCCACCAAGCTTCAAGAGTGTGAGGATTCATTTGTACC 1056  
 982 ATTGGAGACCTAAATCGGAGTCCACCAAGCTTCAAGAGTGTGAGGATTCATTTGTACC 1041  
 1057 CAGATTTGGCAAAATTTTACCAAGCATTTTCAAGGATTAGTATTATATCTATGAAAGCTGTAA 1116



```
Db 1042 AAGATACCTCATTTACCGGCAATTTATATAATATATATATCCGTTATGGTCTCTGAAG 1101
Qy 1117 TAAACTTGGTGAAGGACACAGGTACTATCATTTGAAAACCTTGACAAATGGTCTCTTCC 1176
Db 1102 TAAACTGGTGAAGGACACAGGTCTCTGTTGAAAACACTGGCACTGGAACTCTCGC 1161
Qy 1177 ATTACACCTCTCTTATATTTTA 1198
Db 1162 CTGGATCTGTCTCCATAATA 1183

RESULT 8
US-10-790-589-1
; Sequence 1, Application US/10790589
; Publication No. US20040142376A1
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Krieser, Ronald J.
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCE: DC-0154
; CURRENT APPLICATION NUMBER: US/10790,589
; CURRENT FILING DATE: 2004-03-01
; PRIOR APPLICATION NUMBER: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1224
; TYPE: DNA
; ORGANISM: Mus sp.
US-10-790-589-1

Query Match 53.1%; Score 672.8; DB 17; Length 1224;
Best Local Similarity 75.0%; Pred. No. 1.8e-192;
Matches 871; Conservative 0; Mismatches 282; Indels 9; Gaps 2;

Qy 37 AAACAGAAATGATGGCAAGCTGCTAAGAACATCCTTTGCTTTGCTTCTTCTTGGCTC 96
Db 31 AGATAGAAATGACAGAAAGCTCTAAGAACAGTCTTTCTTTGCTTCTTCTTGGCTC 90
Qy 97 TTTGGGTGCTGGGGCAGCAACAATTTTCATGAGAAATGAAGAGGAAAGCTGTGGAC 156
Db 91 TCTGGGTCTGGGGACACCAGAAATCTCATGCAAAATGAATATGCTGAGCTGTGGAC 150
Qy 157 TGGTTTACTTTTATAGTTTACCTTAAAGACAAAACAAGAAAGTGGAGACTGGGTTA 216
Db 151 TGGTTTATCTTTTATAGTTTACCTTAAAGAGACTAGCAAGGCAAGTGAAGGGGGGCTG 210
Qy 217 GAGTACTCTTACTGACTCTCAACTAGACTGGAGCTGGAGAGAGTGAACAATAATGAAT 276
Db 211 CAGTACTCTTACTGGACTTCCACAGAAACCTGGAAACAGAGCTCTTACCTGATTAAC 270
Qy 277 GACACCAAGAGTGTTTTGGGAAGGACATTAACAAGCTATATACAGCATATGCTCTTAAG 336
Db 271 AGCACCAAGAGTGTCTGGGAGGACCTTACAGCATCTGTATGACACATATTCACG 330
Qy 337 AGTAACAACAGAGCTATCTATATATACATATGATGGAGTCCCTTAAACCTGTGAATTACGT 396
Db 331 AATG---ACACAGCTTATCTAATATACACAGTGTGCTCCCTGGATCTGTGAATTACAGC 387
Qy 397 AGAAGATATGGACACACCAAGGTTTACTGCTGTGGAACAGAGTTCAGGGTCTGGCTG 456
Db 388 AGACAGTATGGACATGCCAAGGTTCTGCTGGTATGGAACAGACAGGAGGTTCTGGCTG 447
Qy 457 ATTCATTCATCCCTCAGTTTCTCTCAATTCGGAAGAGGCTATGATTATCCACCCACA 516
Db 448 ATACACTCTGTTCCCAAGTTTCCCCAGTTC-----ATGGCTATAGTATCCCAACCTCG 501
Qy 517 GGGACAGAAATGGACAAAGTGGCATCTGCATAAATTTTCAAGTACAAACCAAGTATGAGCA 576
```

```
Db 502 GGGAGCGATATGAGCAAAACCGGCAATCTGCATCATCTTCGGATACAGCCAGTTTGAGGA 561
Qy 577 ATAGATTCAGCTCTTGGTCTGCAACCCCAACGCTATAGCTGTCTCCATCCAGCCACC 636
Db 562 ATAGATTTTCACTCTTGGTCTTACAAACCAACATCTACAGCTGCTTCAITCCAGCACC 621
Qy 637 TTTCCACAGAGGCTCATTCACATGCCCGAGCTGTGCAACAGGGCAGCTCATACAGATT 696
Db 622 TTTCACTGGAAACTTATCTACATGCCCGGAGTGTGCCAACTCCAGTTCCTTAAAGATC 681
Qy 697 CCTGGCAGGCTCTCTCACACACTTCAGTCCGCCCCAGGACAAAATTCCTCCATTTGCA 756
Db 682 CCTGTCGGTACCTTCGTGAACTGCATCAGCCCGAGGCTTAAACTTCGTCCATTTGCA 741
Qy 757 AAGTCGATCTCTTTTCTTGACGACATCTTTGACGCTGGATGGCTCAAGGCTCAAGACA 816
Db 742 AAATCAAGTTTATATCTGATGACATCTTTACAGATGGATAGCTCAAAGTTGAAGACA 801
Qy 817 CACTTGTAAAGAACTCTGGCAGGAAAGACAGAGCTTCTTCAAACTGCTCCCTT 876
Db 802 CATTTGTTAGCACAAACCTGGCAGAAAGAAACAGAGCTTCTTCAAACTGTTCCCTG 861
Qy 877 CTTTACCATGCTTACATATATAAAGCAATTAATTTATCAGACACTCTTATTCAGTTCT 936
Db 862 CTTTACCATGCTTACAAACATCAAGTCCATTTGGGGTAACTTCCAGTCTTACAGTTCT 921
Qy 937 TATCAAGATCACGCCCAAGTGTGTATTTCCAAAAGGGCACCACAAAATTCCTGGACATGT 996
Db 922 CGCCAAAGACCATTTCCAAATGTGTGTTTCCATAAAGGGCTCCGCAATTCCTGGACCTGC 981
Qy 997 ATTGGAGACCTAAATCGGAGTCCACCAAGCCTTCAAGAGTGGAGGATTCAITTTGATCC 1056
Db 982 ATTGGAGACCTAAATCGAAGCCTTACCAAGCCTTCAAGAGTGGAGGATTCAITCTGTACA 1041
Qy 1057 CAGAATTTGGCAAAATTTACCAAGCATTTTCAAGGATTTAGTATTTATATATGAAAGCTGTAA 1116
Db 1042 AAGAACTACTACATTTTACCGAGCTTTCATAAATATATCTCCGTTATGGTCTTCTTCC 1101
Qy 1117 TAAACTTGGTGAAGGACACAGGCTACTATCATTTGAAAACCTTGACAAATGGTCTTCTTCC 1176
Db 1102 TAAACTCGGTGAAGGCCACACCCCTCTGCTCTTGAAGAACTGGCACTGGAACTCTGCG 1161
Qy 1177 ATTACACCTCTTATATTTTA 1198
Db 1162 CTGGATCTGTCTCCATAATA 1183
```

## RESULT 9

```
US-09-918-995-7149
; Sequence 7149, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7149
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(480)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-7149
```

Query Match 24.5%; Score 310.8; DB 10; Length 480;  
Best Local Similarity 99.4%; Pred. No. 4.6e-83;  
Matches 312; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 335 AGAGTAAACACACAGCTTCTATATACAAATGATGAGTCCCTAAACCTGTGAATTACA 394  
Db 166 AGAGTAAACACACAGCTTCTATATACAAATGATGAGTCCCTAAACCTGTGAATTACA 225  
QY 395 GTAGAAAGTATGGACACACAAAGGTTTACTGTGTGGAACAGAGTTCAAGGTTCTGGC 454  
Db 226 GCAGAAAGTATGGACACACAAAGGTTTACTGTGTGGAACAGAGTTCAAGGTTCTGGC 285  
QY 455 TGAATTCATCCCTCAGTTTCTCCAAATCCGGAAGAGCTATGATATCCACCCA 514  
Db 286 TGAATTCATCCCTCAGTTTCTCCAAATCCGGAAGAGGCTATGATATCCACCCA 345  
QY 515 CAGGAGACCAAAATGGACAAAGTGGCATCTGCATPACTTTCAAGTACAAACAGATATGAG 574  
Db 346 CAGGAGACCAAAATGGACAAAGTGGCATCTGCATPACTTTCAAGTACAAACAGATATGAG 405  
QY 575 CAATAGATTTCTAGCTCTTGGTCTGCAACCCCAAGCTATAGCTGTCTCCATCCACCCA 634  
Db 406 CAATAGATTTCTAGCTCTTGGTCTGCAACCCCAAGCTATAGCTGTCTCCATCCACCCA 465  
QY 635 CTTTCCACAGGAG 648  
Db 466 CTTTCCACAGGAG 479

RESULT 10

US-10-408-167A-1  
; Sequence 1, Application US/10408167A  
; Publication No. US20030219428A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
Baron, Will F.  
TITLE OF INVENTION: Human DNase II  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/408,167A  
FILING DATE: 04-Apr-2003  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/861034  
FILING DATE: 18-May-2001  
APPLICATION NUMBER: 08/639294  
FILING DATE: 25-Apr-1996

ATTORNEY/AGENT INFORMATION:

NAME: Evans, David W.  
REGISTRATION NUMBER: NONE  
REFERENCE/DOCKET NUMBER: P1024D1C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-1739  
TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1575 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: DNA

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-10-408-167A-1

Query Match 6.5%; Score 82.4; DB 16; Length 1575;  
Best Local Similarity 47.4%; Pred. No. 1.6e-13;  
Matches 470; Conservative 0; Mismatches 486; Indels 36; Gaps 6;

QY 142 GCGAAGCTGTGACCTGGTTTACTTTTATAAGTT---ACCTAAAAGACAAAAAAGAGAA 198  
Db 155 GGCAGCTGTAGACTTGGTTCGTGGTCTACAAGCTGCCAGCTCTTAGAGGTCGGGGAG 214  
QY 199 AGTGGAGAGACTGGTTAGAGTACCTGTACTAGTACTCTCAACTAGAGCTGGAGGAAG 258  
Db 215 GCGGCGCAGAGAGGGCTGCAGTACAAGTATCTGACGAGAGCTCCGGAGGCTCGCGGAC 274  
QY 259 AGTGAGCACTAATGATGACACCAAGAGTGTTTGGGAAGGACATTAACAACAGCTATAT 318  
Db 275 GGCAGGGCCTCATCAACACCGCGGGGGCGGTGGCGAAGCTTCGACCGCTGTAC 334  
QY 319 GAAGCATATGCTTAAAGAGTAACACACAGCCCTATCTAATATATAAATGATGGAGTCCCT 378  
Db 335 CGG-----AGCAACACAGCCAGCTCGCTTCTGTCTTACAAATGACCAACCGCCT 385  
QY 379 AAACC-----TGTGAATTACAGTAGAAAGTATGGACACACCAAGCTTTACTGTGTGG 432  
Db 386 CAACCCAGCAAGGCTCAGGACTCTTCCATGCGTGGGCACACGAAGGTTGCTCTCTCTT 445  
QY 433 AACAGAGTTCAAGGGTTCTGCTGATTCATTCCATCCTCAGTTTCTCCTCAATTCGGAA 492  
Db 446 GACCACGATGGGGCTTCTGGCTGTCCACAGTGTACTTAATCTTCCCTCCACCGGCTCC 505  
QY 493 ---GAAGGCTATGATTATCCACCCACAGGGAGACGAATAATGACAAAGTGGCATCTGATA 549  
Db 506 TCTCTGCATACAGCTGSCCTCATAGCGCTGTACTACGGGCGAGACCTCTCTGTGTG 565  
QY 550 ACTTTCAAGTACACACAGTATGAGCAATAGATTCTCAGCTTTGGTCTGCAACCCCAAC 609  
Db 566 TCTTTTCCCTTCGCTCAGTTCTCGAAGATGGGCAAGCAGCTGACCTACCTACCCCTGG 625  
QY 610 GTCTATAGCTGCTCCATCCAGCCACCTTTTCCACAGGAGCTCANTCAATGCCCGAGCTG 669  
Db 626 GTCTATTAACCTACAGCTGGAAGGGATCTTTGCCAGGAAATTCCTCCGACTTGGAGATGTG 685  
QY 670 TGCACAGGGCCAGCTCATCAGAGATTCTCGGCGGGTCTCTCAACAACCTTCAGTCCGCC 729  
Db 686 GTCAAGGGCCACACAGCTTAGCCAAGAACCTTGAACAGCAGCATCACACTCACATCCCAG 745  
QY 730 CAGGACAAAAATTCCTCCATTTTGCAGAGTCGGATTCTTTTCTTGACGACATCTTTGCA 789  
Db 746 GCGGGGCTGTTTTCAGAGCTTTTCCAAAGTTTCCAAATTTGGAGATGACCTGTACTCC 805  
QY 790 GCCTGGATGGCTCAACGGCTGAAGACACACTTGTTAACAGAAACCTGGCAGCGAAAAAGA 849  
Db 806 GGTGTTGGCAGCAGCCCTTGGTACCAACCTGAGGTCCAGTTCTGGCACAACACTGTA 865  
QY 850 CAAGAGCTTCTTTCAAACTGCTCCCTTCCCTTACCAATGTCTACAATATAAAGCAATAAA 909  
Db 866 GGCATCTCTGCTCTAACTGCTCGATATCTGGCAGGTTCTGATGTGAACCCAGATAGCT 925  
QY 910 TTATCAACGACA-----CTCTTATTTTCACTTATCAGATCAGCCCAAGTGGTATT 963  
Db 926 TTCCCTGGACACAGCGCGCCCAAGCTTCAACAGCAGAGGACCATCTCCAAATGGTGGTG 985  
QY 964 TCCCAAAAGGGCACCACAAAATTCGTGGACATGATTGGAGACTTAATTCGGAGTCCACAC 1023  
Db 986 TCCCAAAAGGGCC-----CTGGACCTCGTGGGTGACATGAATCGGAACACAGGA 1036  
QY 1024 CAAGCTTTCAAGAGTGGAGATTCAATTTGTACCCAGAAATGGCAAAATTTACCAAGCAAT 1083  
Db 1037 GAGGAGCAACGGGGTGGGGGCACACTGTGTGCCAGCTCCAGCCCTCTGGAAGGCTTC 1096  
QY 1084 CAAGGATTTAGTATTATTACTATGAAAGCTGTAA 1115

Db 1097 CAGCGCTGTGAAGAACTACCAGCCCTGTAA 1128

RESULT 11

US-09-908-975-13222

; Sequence 13222, Application US/09908975

; Publication No. US20030165843A1

; GENERAL INFORMATION:

; APPLICANT: SHOSHAN, Avi

; APPLICANT: WASSERMAN, Alon

; APPLICANT: MINTZ, Eli

; APPLICANT: MINTZ, Liat

; APPLICANT: FAIGLER, Simchon

; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICING

; FILE REFERENCE: 36688-0005

; CURRENT APPLICATION NUMBER: US/09/908,975

; CURRENT FILING DATE: 2001-07-20

; PRIOR APPLICATION NUMBER: US 60/287,724

; PRIOR FILING DATE: 2001-05-02

; PRIOR APPLICATION NUMBER: US 60/221,607

; PRIOR FILING DATE: 2000-07-28

; NUMBER OF SEQ ID NOS: 32337

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 13222

; LENGTH: 60

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-908-975-13222

Query Match 4.7%; Score 60; DB 10; Length 60;

Best Local Similarity 100.0%; Pred. No. 1.1e-07;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 986 GCTGGACATGTTATGGAGACCTAAATCGGAGTCCACACCAAGCTTCAGAGTGGAGAT 1045

Db 1 GCTGGACATGTTATGGAGACCTAAATCGGAGTCCACACCAAGCTTCAGAGTGGAGAT 60

RESULT 12

US-09-908-975-2418

; Sequence 2418, Application US/09908975

; Publication No. US20030165843A1

; GENERAL INFORMATION:

; APPLICANT: SHOSHAN, Avi

; APPLICANT: WASSERMAN, Alon

; APPLICANT: MINTZ, Eli

; APPLICANT: MINTZ, Liat

; APPLICANT: FAIGLER, Simchon

; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICING

; FILE REFERENCE: 36688-0005

; CURRENT APPLICATION NUMBER: US/09/908,975

; CURRENT FILING DATE: 2001-07-20

; PRIOR APPLICATION NUMBER: US 60/287,724

; PRIOR FILING DATE: 2001-05-02

; PRIOR APPLICATION NUMBER: US 60/221,607

; PRIOR FILING DATE: 2000-07-28

; NUMBER OF SEQ ID NOS: 32337

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 2418

; LENGTH: 65

; TYPE: DNA

; ORGANISM: Rattus norvegicus

US-09-908-975-2418

Query Match 4.5%; Score 57; DB 10; Length 65;

Best Local Similarity 92.3%; Pred. No. 9.3e-07;

Matches 60; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 984 TCGCTGGACATGTTATGGAGACCTAAATCGGAGTCCACACCAAGCTTCAGAGTGGAGG 1043

Db 1 TCGCTGGACCTGTTATGGAGACCTAAATCGGAGTCCACACCAAGCTTCAGAGTGGAGG 60

QY 1044 ATTCA 1048

Db 61 ATTTA 65

RESULT 13

US-09-918-995-1186

; Sequence 1186, Application US/09918995

; Publication No. US20030073623A1

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED FROM VARIOUS CDNA LIBRARIES

; FILE REFERENCE: 20411-756

; CURRENT APPLICATION NUMBER: US/09/918,995

; CURRENT FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: US/09/235,076

; PRIOR FILING DATE: 1999-01-20

; NUMBER OF SEQ ID NOS: 38054

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1186

; LENGTH: 478

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)..(478)

; OTHER INFORMATION: n = A,T,C or G

US-09-918-995-1186

Query Match 4.0%; Score 50.6; DB 10; Length 478;

Best Local Similarity 45.5%; Pred. No. 0.00031;

Matches 179; Conservative 0; Mismatches 214; Indels 0; Gaps 0;

QY 527 ATGACACAAAGTGGCATCTGCATTAACCTTTCAAGTACACCAAGTATGAGGCAATGATTCTC 586

Db 35 ACGGGCAGACCTCTCTGTGTCTCTTTCCCTTCGCTCAGTCTCTGAAGATGGGCAAGC 94

QY 587 AGCTCTGGTCTGCAACCCCAACGCTCTATAGCTGTCTCCATCCCAAGCCACCTTTTCCAGG 646

Db 95 AGCTGACCTACACTACCTACCTCCCTGGGTCTATACTACCAAGTGGAGGATCTTTGCCAGG 154

QY 647 AGCTCATTCACATGCCCCAGCTGTGACACAGGCCAGCTCATCAGAGTTCTCTGGCAGGC 706

Db 155 AATCCCCGACTTGGAGAATGTGTCAGGGCCACCAAGTTCAGCAAGAACCTTGAACA 214

QY 707 TCCTACCAACATTCAGTGGCCCGCCAGGACAAAAATTTCTCCATTTTGCAGGCTGGATT 766

Db 215 GCAGCATCAGCTCATCATCCAGCCGGGCTGTTTTCAGAGCTTGGCAAGTTGAGCA 274

QY 767 CTCTTCTTGACGACATCTTTTGGAGCTGGATGGCTCAACGGCTGAGACACACACTTGTAA 826

Db 275 AATTTGGAGATGACCTGTACTCCGGCTGTGTCAGCAGCCCTTGGTACCACTTCAGG 334

QY 827 CAGAACCTTGGAGCGGAGGAGGAGAGCTTCTCTTCAAACTGCTCCCTTCCCTTACCAG 886

Db 335 TCCAGTTCTGGACACAAACTGTAGGCATCTGCGCTCTAACTGCTCGGATATCTGGCAGG 394

QY 887 TCTACAATATAAAGCAATTAAATATATCAGAC 919

Db 395 TTCTGATGTGAACAGATAGCTTTCCCTGGAC 427

RESULT 14

US-09-983-965-4982

; Sequence 4982, Application US/09983965

; Patent No. US20020137160A1

; GENERAL INFORMATION:

; APPLICANT: Warren, Wesley C.

; APPLICANT: Tao, Nengbing

; APPLICANT: Byatt, John C.

; APPLICANT: Mathialagan, Nagappan



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2004, 00:41:52 ; Search time 3092.15 Seconds  
(without alignments)  
12245.607 Million cell updates/sec

Title: US-10-790-589-3

Perfect score: 1268

Sequence: 1 atggggaagtgctcgtg.....tctctcatgtttaccattta 1268

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

EST.*
1: em_estba.*
2: em_esthum.*
3: em_estlin.*
4: em_estnu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_nam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rtd.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	870.4	68.6	893	13	BQ681950
C 2	723.6	57.1	758	14	CD365871
C 3	707.2	55.8	729	14	CD365107
C 4	690.6	54.5	727	14	CD366106

5	681.2	53.7	923	12	BG761484
6	672.2	53.0	2689	11	AK079224
C 7	621.8	49.0	676	9	AI802154
C 8	580.4	45.8	609	12	BG149286
C 9	553	43.6	554	9	AI420898
C 10	550.2	43.4	571	10	AW295184
C 11	528	41.6	536	9	AI694455
C 12	511.4	40.3	521	9	AI820599
C 13	482.8	38.1	487	9	AI791998
C 14	467.8	36.9	482	13	BX380577
C 15	441.8	34.8	446	9	AA988125
C 16	415	32.7	448	10	BF108594
C 17	395.6	31.2	714	10	BE966684
C 18	349.2	27.5	397	13	BX482994
C 19	315	24.8	695	14	CF725003
C 20	303.4	23.9	521	13	BQ373395
C 21	282.4	22.3	887	13	BU241430
C 22	280.2	22.1	510	9	AI526873
C 23	272.4	21.5	710	13	BU240964
C 24	256.4	20.2	546	13	BX526247
C 25	227.6	17.9	432	9	AI875954
C 26	225.2	17.8	792	13	BU240892
C 27	220.8	17.4	669	12	BI392355
C 28	216	17.0	755	13	BU397846
C 29	210	16.6	258	10	BE937933
C 30	209	16.5	800	13	BU457064
C 31	207.6	16.4	543	9	AI048641
C 32	207	16.3	999	13	BU235150
C 33	203.6	16.1	643	13	BU215085
C 34	202.2	15.9	406	14	CB808400
C 35	198.2	15.6	522	13	BU299065
C 36	194	15.3	424	14	CB796507
C 37	191	15.1	473	14	CB729923
C 38	188	14.8	381	9	AI876076
C 39	187.8	14.8	223	9	AI658810
C 40	186.4	14.7	652	10	BB627834
C 41	186.2	14.7	643	13	BU280296
C 42	182	14.4	948	14	CF590064
C 43	160.4	12.6	565	12	BM426634
C 44	150.6	11.9	476	9	AI549457
C 45	146	11.5	454	12	BI281397

#### ALIGNMENTS

RESULT 1  
BQ681950  
LOCUS  
DEFINITION BQ681950 893 bp mRNA linear EST 15-JUL-2002  
5', mRNA sequence. AGENCOURT\_818335 NIH\_MGC\_112 Homo sapiens cDNA clone IMAGE:6262400  
ACCESSION BQ681950  
VERSION BQ681950.1 GI:21794629  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 893)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
Tissue Procurement: DCTD/DTF  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILLNL at: <http://image.llnl.gov>  
Plate: LLCN2424 row: k column: 09  
High quality sequence stop: 647.

BG761484 602718694  
AK079224 Mus muscu  
AI802154 tx30d07.x  
BG149286 nad2sg11.  
AI420898 tf03g06.x  
AW295184 UI-H-B12-  
AI694455 wd83h10.x  
AI820599 os07c09.y  
AI791998 os07c09.y  
BX380577 BX280577  
AA988125 os07c09.s  
BF108594 Th60f04.x  
BE966684 601661373  
BX482994 DXF2p686B  
CF725003 UI-M-GZ0-  
BQ373395 QV1-FT020  
BU241430 603323735  
AI526873 uj47d10.x  
BU240964 603322469  
BX526247 BX526247  
AI875954 uj53b10.y  
BU240892 603322279  
BI392355 pppin.pk0  
BU397846 603533958  
BE937933 CM0-TN003  
BU457064 603770721  
AI048641 ud64c11.y  
BU235150 603791183  
BU215085 603750777  
CB808400 AMGNNUC:S  
BU299065 603734691  
CB796507 AMGNNUC:S  
CB729923 AMGNNUC:S  
AI876076 uj55b03.y  
AI658810 ta08h10.x  
BB627834 BB627834  
BU280296 603863307  
CF590064 AGENCOURT  
BM426634 p9f2n.pk0  
AI549457 UI-R-C3-t  
BI281397 UI-R-DC0-

Db	721	CAGGATTAGTATTATATGTAAGAGCTGTAAAGTAAACTTGGTGAAGAGCACAGGTACTA	780
QY	1145	TCATTGAAAACTTGACAACTGGCTCTTCTTCATTACACCTTCTTTATATTTTAAAGGCC	1204
Db	781	TCATTGAAAACTTGACAACTGGCTCTTCTTCATTACACCTTCTTTATATTTAAGGGCC	840
QY	1205	TGTGAATATACCTTAAACCTGCATATACAAATAAAACATATTTCTCTC	1254
Db	841	TGTGGATATACCTTAAACCTGCATATACAAATAAAACATATTTCTCTC	890
RESULT 2			
CD365871/c			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
QY	365	ATGATGGAGTCCCTAAACCTGTGAATTACAGTAGAAGATATGACACACCAAGGTTTAC	424
Db	1	ATGATGGAGTCCCTAAACCTGTGAATTACAGTAGAAGATATGACACACCAAGGTTTAC	60
QY	425	TGCTGTGGAAACAGAGTTCAAGGGTTCTGGCTGATTTCATTCATCCATCCTCAGTTTCTCTCAA	484
Db	61	TGCTGTGGAAACAGAGTTCAAGGGTTCTGGCTGATTTCATTCATCCATCCTCAGTTTCTCTCAA	120
QY	485	TTCCGGAAGAAGCGTAGATTATTCACCCACAGGAGACGAATATGACCAAAAGTGGCATCT	544
Db	121	TTCCGGAAGAAGCGTAGATTATTCACCCACAGGAGACGAATATGACCAAAAGTGGCATCT	180
QY	545	GCATAACTTTCAAGTACAAACAGATATGAGCAATAGATTCTCAGCTCTTGGTCTGCAACC	604
Db	181	GCATAACTTTCAAGTACAAACAGATATGAGCAATAGATTCTCAGCTCTTGGTCTGCAACC	240
QY	605	CCACGCTATAGTGTCTCCATCCAGCCACCTTTCCACAGAGCTCATTCATGCGCCC	664
Db	241	CCACGCTATAGTGTCTCCATCCAGCCACCTTTCCACAGAGCTCATTCATGCGCCC	300
QY	665	AGCTGTGCACACAGGCGCAGCTCATCAGAGATTCCTGGCAGGCTCTTCCACCACTTCAGT	724
Db	301	AGCTGTGCACACAGGCGCAGCTCATCAGAGATTCCTGGCAGGCTCTTCCACCACTTCAGT	360
QY	725	CGGCCACAGGACAAAATTTCTTCATTTGCAAGTGGATCTTTTCTTGAGCATCT	784
Db	361	CGGCCACAGGACAAAATTTCTTCATTTGCAAGTGGATCTTTTCTTGAGCATCT	420
QY	785	TTGCAGCTCGATGGCTCAACGGCTGAAGACACACTTTGTTAAACAGAACTGGCAGCGAA	844
Db	421	TTGCAGCTCGATGGCTCAACGGCTGAAGACACACTTTGTTAAACAGAACTGGCAGCGAA	480
QY	845	AAAGACAAGAGCTTCTTCAAACTGCTCCCTTCCATTCAGTGTCTCAATATATAAGCAA	904
Db	481	AAAGACAAGAGCTTCTTCAAACTGCTCCCTTCCATTCAGTGTCTCAATATATAAGCAA	540
QY	905	TTAAATTTATCAGCACTCTTATTTTCAAGTTCTTTATCAAGATCATCGCAAGTGTGATTT	964
Db	541	TTAAATTTATCAGCACTCTTATTTTCAAGTTCTTTATCAAGATCATCGCAAGTGTGATTT	600
QY	965	CCCAAAAGGGCACCAAAATTCGCTGGACATGTATTTGGAGACCTTAATTCGAGTCCACACC	1024
Db	601	CCCAAAAGGGCACCAAAATTCGCTGGACATGTATTTGGAGACCTTAATTCGAGTCCACACC	660
QY	1025	AAGCCCTTCAGAGTGGAGATTCATTTGTATCCAGAAATTCGCAAAATTTTCAAGCATTTT	1084
Db	661	AAGCCCTTCAGAGTGGAGATTCATTTGTATCCAGAAATTCGCAAAATTTTCAAGCATTTT	720
QY	1085	AAGGATTAGTATTATATGTAAGTAACTTGGTGAAGAGCACAGGTACTA	1144



```

RESULT 4
CD366106/c
LOCUS
DEFINITION
  UI-H-FT1-bjs-1-13-0-UI.s1 NCI CGAP_FTI Homo sapiens cDNA clone
  UI-H-FT1-bjs-1-13-0-UI 3', mRNA sequence.
ACCESSION
  CD366106
VERSION
  CD366106.1 GI:31150196
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 727)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgabs-remail.nih.gov
  Tissue Procurement: Dr. Gary W. Humminghake, U of I
  CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
  DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
  CDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: Distribution information can be found at
  http://genome.uiowa.edu/distribution/cgap.html
  Seq primer: M13 FORWARD
  POLYA=Yes.
FEATURES
  source
  1..727
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="UI-H-FT1-bjs-1-13-0-UI"
  /tissue_type="Aveolar Macrophage"
  /dev_stage="Adult"
  /lab_host="DH10B (Life Technologies)"
  /clone_lib="NCI-CGAP_FTI"
  /note="Origin: Lung; Vector: pVT73-Pac (Pharmacia) with a
  modified polylinker; Site 1: EcoR I; Site 2: Not I;
  NCI CGAP FTI is a normalized cDNA library constructed from
  a pool of 81 RNA samples from Alveolar Macrophages
  challenged with different treatments. The library was
  normalized according to Bonaldo, Lennon and Soares, Genome
  Research, 6:791-806, 1996. First strand cDNA synthesis was
  primed with an oligo-dT primer containing a Not I site.
  Double stranded cDNA was ligated to an EcoR I adaptor,
  digested with Not I, and cloned directionally into
  pVT73-Pac vector. The oligonucleotide used to prime the
  synthesis of first-strand cDNA contains a library tag
  sequence that is located between the Not I site and the
  (dT)18 tail. The sequence tag for this library is
  GGCCATGCGG. The tissue was provided by Dr. Gary W.
  Humminghake of the University of Iowa.
  TAG TISSUE=Human Lung Aveolar Macrophage
  TAG LIB=UI-H-FT1
  TAG_SEQ=GGCCATGCGG"
ORIGIN
  Query Match 54.5%; Score 690.6; DB 14; Length 727;
  Best Local Similarity 99.0%; Pred. No. 1.4e-182;
  Matches 704; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
  QY 558 GTACAAACAGTAGGCAATGAGATTCAGCTCTTGCTGCAACCCCAAGCTCTATAG 617
  DB 727 GTACACACAGTAGA-GCAATAGATTCAGCTCTTGCTGCAACCCCAAGCTCTATAG 669
  QY 618 CTGCTCCATCCGACCACTTCCACAGGAGTCTATTACATGCCCGCCAGCTGGACACG 677
  DB 668 CTGCTCCATCCGACCACTTCCACAGGAGTCTATTACATGCCCGCCAGCTGGACACG 609
  QY 678 GGCCAGCTCATCAGAGATTCCTGGCAGGCTCCTCACCACATTCAGTCGCCCGCAGGACA 737

```

```

Db 608 GCCAGNCTCATCAGAGATTCCTGGCAGGCTCCTCACCACATTCAGTCGCCCGCAGGACA 549
QY 738 AAAATTCCTCATTTTGCAGAGTCGATTTCTTTCTTGACGACATCTTTTCAGCTCGAT 797
Db 548 AAAATTCCTCATTTTGCAGAGTCGATTTCTTTCTTGACGACATCTTTTCAGCTCGAT 489
QY 798 GGCTCAACGGCTGAAGACACATTTGTTAAACAGAAACCTGGCAGCGAAACAGCAAGAGCT 857
Db 488 GGCTCAACGGCTGAAGACACATTTGTTAAACAGAAACCTGGCAGCGAAACAGCAAGAGCT 429
QY 858 TCCTTCAAACTGCTCCCTTCTTACCATGCTCAATATATAAAGCAATTAATATACAG 917
Db 428 TCCTTCAAACTGCTCCCTTCTTACCATGCTCAATATATAAAGCAATTAATATACAG 369
QY 918 ACATCTTATTTAGTTTCTTATCAAGATCAGCGCAAGTGGTGTATTTCCCAAAAGGCGAC 977
Db 368 ACATCTTATTTAGTTTCTTATCAAGATCAGCGCAAGTGGTGTATTTCCCAAAAGGCGAC 309
QY 978 CAAAAATCGCTGGACATGATTTGGAGACCTAAATCGGAGTCCACACCAAGCCTTCAAGAG 1037
Db 308 CAAAAATCGCTGGACATGATTTGGAGACCTAAATCGGAGTCCACACCAAGCCTTCAAGAG 249
QY 1038 TGGAGGATTCATTTGTATCCCGAATTTGGCAAAATTTACAGCATTTCAAGATTTAGTATT 1097
Db 248 TGGAGGATTCATTTGTATCCCGAATTTGGCAAAATTTACAGCATTTCAAGATTTAGTATT 189
QY 1098 ATACTATGAAGCTGTAAAGTAAATCTGGTGAAGGACACAGCTACTCATTTGAAAACCT 1157
Db 188 ATACTATGAAGCTGTAAAGTAAATCTGGTGAAGGACACAGCTACTCATTTGAAAACCT 129
QY 1158 TGACAAATGGGTCTTCTTCCATTTACACCTTCTTTATATTTTAAAGCCTGTGAATATCTT 1217
Db 128 TGACAAATGGGTCTTCTTCCATTTACACCTTCTTTATATTTTAAAGCCTGTGAATATCTT 69
QY 1218 ATAACTGTCATATCAGAAATATAACATATTTCTCTCATGTTTACCATTTA 1268
Db 68 ATAACTGTCATATCAGAAATATAACATATTTCTCTCATGTTTACCATTTA 18
RESULT 5
BG761484 923 bp mRNA linear EST 15-MAY-2001
602718694F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4858894 5',
mRNA sequence.
BG761484
BG761484.1 GI:14072137
EST.
KEYWORDS
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 923)
  NIH-MGC http://mgc.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgabs-remail.nih.gov
  Tissue Procurement: ATCC/DCTD/DTF
  CDNA Library Preparation: Ling Hong/Rubin Laboratory
  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Incyte Genomics, Inc.
  Clone Distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  Plate: LLCM1713 row: k column: 23
  High quality sequence stop: 845.
  Location/Qualifiers
  1..923
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="IMAGE:4858894"
FEATURES
  source

```



```

/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 43"
/notes="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using Zap-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

```

## ORIGIN

Query Match	53.7%;	Score 681.2;	DB 12;	Length 923;
Best Local Similarity	95.9%;	Pred. No. 6.3e-180;		
Matches 821;	Conservative 0;	Mismatches 13;	Indels 22;	Gaps 11;
QY	367	GATGGAGTCCTAAACCTGTGAATTACAGTAGAAAGATGACACACACCAAGGTTTACTG	426	
DB	1	GATGGAGTCCTAAACCTGTGAATTACAGCAGAAAGTATGGACACACCAAGGTTTACTG	60	
QY	427	CTGTGGAACAGAGTTCAAGGGTTCTGGCTGATTCATTCATCCCTCAGTTTCCTCCAAATT	486	
DB	61	CTGTGGAACAGAGTTCAAGGGTTCTGGCTGATTCATTCCTCATCCCTCAGTTTCCTCCAAATT	120	
QY	487	CCGGAAGAAGGCTATGATTAATCCACCACAGGAGAGCAGAAATGGACAAAGTGGCATCTGC	546	
DB	121	CCGGAAGAAGGCTATGATTAATCCACCACAGGAGAGCAGAAATGGACAAAGTGGCATCTGC	180	
QY	547	ATRACTTTCAAGTACAAACCAATGATGAGCAATAG-----ATTCTCAGCTCTTTGG	595	
DB	181	ATAACTTTCAAGTACAAACCAATGATGAGCAATAGGTAATAAAATCTCAGCTCTTTGG	240	
QY	596	TCGCAACCC--AACGCTATAGCTGCTCATCC--AGCCACCTTTCCACAGAGCTCAT	653	
DB	241	TCGCAACCCCTAACGCTATAGCTGCTCATCCCTAGCCACCTTTCCACAGAGCTCAT	300	
QY	654	TCAATGCCCCA--GCTGTGCA--CCAGGGCCAGCTCATCAGAGATTCCTGCGAGGCTCCTC	711	
DB	301	TCATATGCCCATGCTGTGCAGCCAGGGCCAGCTCATCAGAGATTCCTGCGAGGCTCCTC	360	
QY	712	ACCACACTTCAGTCGGCCGACGGACAAAATTCCTCCATTTTGGAAAGTCGATCTTTTT	771	
DB	361	ACCACACTTCAGTCGGCCGACGGACAAAATTCCTCCATTTTGGAAAGTCGATCTTTTT	420	
QY	772	CTTGACGACATCTTTGCAGCTCGATGGCTCAACGGCTGAAGACACACTTTGTTAAACAGAA	831	
DB	421	CTTGACGACATCTTTGCAGCTCGATGGCTCAACGGCTGAAGACACACTTTGTTAAACAGAA	480	
QY	832	ACCTGGC--AGCGAAAAACAGACGCTTCCTTCAAACGTCTCCCTTCCTTACCATGTCT	889	
DB	481	ACCTGGCTAGCGAATAAAGACAGAGCTTCCTTCAAACGTCTCCCTTCCTTACCATGTCT	540	
QY	890	ACAAATATAAAGCAATTAATATACGACACCTTTATTCAGTTCTTATCAAGATCAG	949	
DB	541	ACAAATATAAAGCAATTAATATACGACACCTTTATTCAGTTCTTATCAAGATCAG	600	
QY	950	CCAAAGTGGTATTTCCCAAAAGGCCACCAA--AAATCGCTGGACATGTATTTGGAGACCTTA	1008	
DB	601	CCAAAGTGGTATTTCCCAAAAGGCCACCAAAGATTCGCTGGACATGTATTTGGAGACCTTA	660	
QY	1009	AATCGAGTCCACACCAAGGCTTCAGAGTGG--AGGATTCATTTGTACCC--AGAAATGGC	1066	
DB	661	AATCGAGTCCACACCAAGGCTTCAGAGTGGAGGATTCATTTGTATCCCAAGATTTGGC	720	
QY	1067	AAATTTTACCAAGCATTTCAAGGATAGTATATATCATGAAAGCTCTTAAGTAAACTTGGT	1126	
DB	721	AAATTTTACCAAGCATTTCAAGGATAGTATATATCATGAAAGCTCTCAATTAACCTTGGT	780	
QY	1127	GAAAGACACAGGTACTAT--CAATGAAAACTTTGACAAATGGGCTTTCTTCCATTAACCT	1185	
DB	781	GAAAGACCCAGGTACTATCCATGAAAACTTTGACAAATGGGCTTTCTTCCATTA--ACCT	839	

QY 1186 TCITTATATTTAAAG 1201  
Db 840 TCITTATATTTGAAG 855

RESULT 6	
AK079224	
LOCUS	
DEFINITION	

LOCUS	AK079224	2689 bp	mRNA	linear	HTC 19-SEP-2003
DEFINITION	Mus musculus adult male urinary bladder cDNA, RIKEN full-length enriched library, clone:5530041E10 product:deoxyribonuclease II beta, full insert sequence.				

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

ORGANISM	Mus musculus
EXTRACT	Eukaryota; Metazoa; Chordata; Craniata; Eukaryota; Metazoa; Chordata; Craniata; Mammalia; Eutheria; Rodentia; Sciurognath
REFERENCE	1
AUTHORS	Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636

REFERENCE	2
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Suganuma, T., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., Tanabe, T., and Tashiro, K.
TITLE	Normalization and subtraction of cap-trapper transcripts from total RNA to prepare full-length cDNA libraries for RNA sequencing
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159

3  
REFERENCE  
AUTHORS  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, K., Kono, H., Akiyama, J., Nishi, K., Kitsuai, Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, Y., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka Okazaki, Y., Muramatsu, M., Inoue, Y., Kira

TITLE	RIKEN integrated sequence analysis (RISA) sequencing pipeline with 384 multicapillary
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861

REFERENCE	4	THE RIKEN GENOME EXPLORATION RESEARCH GROUP PHASE II TEAM AND THE FANTOM CONSORTIUM.
AUTHORS		FUNCTIONAL ANNOTATION OF A FULL-LENGTH MOUSE cDNA COLLECTION
TITLE		NATURE 409, 685-690 (2001)
JOURNAL		

REFERENCE	5 Nature 403, 693-699 (2001)
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)

6 (bases 1 to 2689) 120927

REFERENCE  
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arai, K., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohsumi, N., Okazaki, Y., Saito, R., Saiboh, H., Sakai, C., Sakai, K., Sakazume, N., Sauc, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sugabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akashira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

**TITLE**  
**JOURNAL**  
**Direct Submission**  
**Submitted (16-APR-2002)** Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Parameters, H. and Hayashizaki, I.

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp).  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/.

# FEATURES

Location/Qualifiers  
1. .2689  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="PANTOM\_DB:9530041E10"  
/db\_xref="MGI:2399700"  
/db\_xref="taxon:10090"  
/clone="9530041E10"  
/sex="male"  
/tissue\_type="urinary bladder"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
335. .1399  
/note="unnamed protein product; deoxyribonuclease II beta  
(MGI:1913283, GB|NM\_019957, evidence: BLASTN, 99%,  
match=1529)  
putative"  
/codon\_start=1  
/protein\_id="BAC37579.1"  
/db\_xref="GI:26347861"  
/translation="MTAKPIKTVLSLFFALSGVLGTPETISCRNEYGEAVDFIFYL  
PKTSKSEERAGLYLIDSTQWNSLYLINSALGKTLQHLVDHNSINDTAY  
LIYNDGPGSVNYSQVGHAKGLLWNPETQGFVLIHSPKPPVHGVYPTSGRRYGO  
TGICITGYSOFSEIDFOLLVQENIYSCFPSPFHWKLIYMPKNCANSSSLKIPVY  
LAELHSAQGLNHFHFKSFTYDDIFTGWIQAQKHLHLLAFTWKKQKQBELPNSCLPY  
HVYNIKGSITKSYFSRSDPHSKVCSIKGSANRWTCIGDLNLSLHQAURGGGFICT  
KNHYIQAFHKLVLRYLPGCK"

# ORIGIN

Query Match 53.0%; Score 672.2; DB 11; Length 2689;  
Best Local Similarity 75.0%; Pred. No. 2.9e-177;  
Matches 883; Conservative 0; Mismatches 283; Indels 11; Gaps 3;  
QY 37 AAACAGAAATGATGGCAAGACTGCTAAGAACATCTTTGCTTTGCTCTCTTGGCCTC 96  
DB 326 AGATAGAAATGACAGCAAGGCTTAAGACAGTTCTTTCTTCTCTTCTTGGCCTC 385  
QY 97 TTTGGGTGCTGGGGAGCAACAATTTTCATGCAAGATGAAGAGGGAAGCTGTGGAC 156  
DB 386 TCTGGGTCTCTGGGACACCAAGAAATCTCATGCAAGATGAATATGTTGAAGCTGTGGAC 445  
QY 157 TGGTTTACTTTTATAGTTACTTAAAGACAAACAAAGGAGGAGGAGCTGCTGTTA 216  
DB 446 TGGTTTACTTTTATAGTTACTTAAAGACAAACAAAGGAGGAGGAGGAGCTGTTA 505  
QY 217 GAGTACCTGTACTAGACTCTTACAATCTAGAGCTGGAGGAGGAGGAGCACTAATGAAT 276  
DB 506 CAGTACCTGTACTAGACTCTTACAATCTAGAGCTGGAGGAGGAGGAGCACTAATGAAT 565  
QY 277 GACACAGAGGTTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 336  
DB 566 AGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 625  
QY 337 ACTAACACACAGGCTTCTTATATATATATATATATATATATATATATATATATATAT 396  
DB 626 AATGACA---CAGCCTATCTATATATATATATATATATATATATATATATATATAT 682  
QY 397 AGAAGATATGACACACCAAGGTTTACTGTGCTGTGGAACAGAGTTCAAGGGTTCTGGCTG 456

683 AGACAGTATGACATGCCAAAGGTTCTGCTGTATGGAACAGAAAGCGAGGGTTCTGGTG 742  
QY 457 ATTCAATTCATTCCTCAGTTTCTCCTCCAAATCCGGAAGAGGATGATATATCCACCCACA 516  
DB 743 ATACACTCTCTTCCAAAGTTTCCCCAGTTTCTCCTCCAAATCCGGAAGAGGATGATAT 796  
QY 517 GGAGAGCAATGACAAAGTGGCATCTGCATCACTTTCAAGTACACCAAGGATGAGGCA 576  
DB 797 GGAGGCGATATGACAAACCGGCATCTGCATCACTTTCCGATACAGCCAGTTTGGAA 856  
QY 577 ATAGATTCTCAGCTCTTGGTCTGCAACCCCAACCCCAACCCCAACCCCAACCCCAAC 636  
DB 857 ATAGATTCTCAGCTCTTGGTCTTACAAACCAACATCTACAGCTCTTCAATCCAGCACC 916  
QY 637 TTTTACCAGGAGCTCATTACATGCCCCAGCTGTCACCCAGGCGAGCTCATCAGAGATT 696  
DB 917 TTTTACCAGGAGCTCATTACATGCCCCAGCTGTCACCCAGGCGAGCTCATCAGAGATT 976  
QY 697 CCTGGCAGGCTCTCTCACCACATCTTCACTCGGCCCCAGGAGCAAAAATCTCTCCATTTGCA 756  
DB 977 CCTGTCGGTACCTCGCTGAACTGCACTCAGCCCGAGGCTTAACTTCTGCTCCATTTGCA 1036  
QY 757 AAGTCGATCTTTTCTGAGGACATCTTTGACGCTGGATGGCTCAAGGCTGAGGCA 816  
DB 1037 AAATCAAGTTTATATATGATGATGATCTTTACAGGATGATGATCTCAAAAGTTGAGGCA 1096  
QY 817 CACTTGTATACAGAAACCTGCGAGGAAAGAGCAAGAGCTTCTTCAAACTGCTCCCTT 876  
DB 1097 CATTGTTAGCACAACCTGCGAGGAAAGAGCAAGAGCTTCTTCAAACTGCTCCCTT 1156  
QY 877 CTTTACCAGTCTTCAATATATAAAGCAATTAATATATACGACATCTTTATTTTCAAGTTCT 936  
DB 1157 CTTTACCAGTCTTCAATATATAAAGCAATTAATATATATACGACATCTTTATTTTCAAGTTCT 1216  
QY 937 TATCAAGATCAGCGAGTGTGTATTTCCCAAGGCAACAAATTCGCTGACATGT 996  
DB 1217 CGCCAAAGCAATTCCTTCAATATGTTGTTTCCATAAAGGCTCCGCAATTCGCTGACATGT 1276  
QY 997 ATTGGAGACCTTAAATCGGAGTCCACCAAGCTTTCAAGAGTGGAGGATTCATTTGTACC 1056  
DB 1277 ATTGGAGACCTTAAATCGGAGTCCACCAAGCTTTCAAGAGTGGAGGATTCATTTGTACC 1336  
QY 1057 CAGAAATGCGCAAAATTTACCAAGCATTTCAAGGATTTATATATATATATATATATAT 1116  
DB 1337 AAGATCACTACATTTTACAGGCAATTTTCAAAATTTATATATATATATATATATAT 1396  
QY 1117 TAACTTGGTGAAGGACACAGTACTATCTTGAAGGCTTGAAGGCTTGAAGGCTTCTTCC 1176  
DB 1397 TAACTTGGTGAAGGACACAGTACTATCTTGAAGGCTTGAAGGCTTGAAGGCTTCTTCC 1456  
QY 1177 ATTACACC--TTCTTTATATTTTAAAGGCTGTGAAT 1211  
DB 1457 CTTGATCTGTTCTCCATTAATTTCAAGGCTTCTGAGT 1493

RESULT 7  
AI802154/c  
LOCUS  
DEFINITION  
AI802154  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

AI802154  
676 bp mRNA linear EST 16-DEC-1999  
Cx30d07.x1 NCI CGAP Lu24 Homo sapiens cDNA clone IMAGE:2271085.3,  
similar to TR:090759 090759 DBXNRIBONUCLEASE II ; mRNA sequence.  
AI802154  
GI:5367626  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 676)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Sequencing by: Greg Lennon, Ph.D.

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 982 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 456.

## FEATURES

source

Location/Qualifiers  
1. .676  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2211085"  
/tissue\_type="carcinoid"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Lu24"  
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP Lu5 was prepared, and ss circles were used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clones 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 49.0%; Score 621.8; DB 9; Length 676;  
Best Local Similarity 97.5%; Pred. No. 2.8e-163;  
Matches 663; Conservative 0; Mismatches 13; Indels 4; Gaps 3;  
  
QY 574 GCAATAGATCTCAGCTCTGGTGTGCTGCAACCCCAAGCTCTATAGCTGCTCCATCCAGCC 633  
DB 676 GCAATAGATCTCAGCTCTGGTGTGCTGCAACCCCAAGCTCTATAGCTGCTCCATCCAGCC 617  
QY 634 ACCTTTCCACGAGGCTCATTTCATGCCCCCAGCTGTGCAACGGGCGAGCTCATCAGAG 693  
DB 616 ACCTTTCCACGAGGCTCATTTCATGCCCCCAGCTGTGCAACGGGCGAGCTCATCAGAG 557  
QY 694 ATTCTCTGGCAGGCTCTCTCACCACATTCAGTGGCCCGCCAGGACAAAATTCCTCAATTTT 753  
DB 556 ATTCTCTGGCAGGCTCTCTCACCACATTCAGTGGCCCGCCAGGACAAAATTCCTCAATTTT 500  
QY 754 GCAAGTCGGATCTTTTCTTGACGACATCTTTGCGAGCTGGATGGCTCAACGGCTGAAG 813  
DB 499 GCAAGTCGGATCTTTTCTTGATGACATCTTTGCGAGCTGGATGGCTCAACGGCTGAAG 440  
QY 814 ACACACTGTGTTAACAAACCTGGCGAGCGAAGAACAGACAGCTTCTTCAAACTGCTCC 873  
DB 439 ACACACTGTGTTAACAAACCTGGCGAGCGAAGAACAGACAGCTTCTTCAAACTGCTCC 380  
QY 874 CTTCTCTACCACTGTCTACAAATATAAAGCAATATAATATACAGACACTCTTATTTTCAGT 933  
DB 379 CTTCTCTACCACTGTCTACAAATATAAAGCAATATAATATACAGACACTCTTATTTTCAGT 320  
QY 934 TCTTATCAAGATCAGCCCAAGTGGTGTATTTCCAAAAGGGCACCAAAAATCGCTGGACA 993  
DB 319 TCTTATCAAGATCAGCCCAAGTGGTGTATTTCCAAAAGGGCACCAAAAATCGCTGGACA 260  
QY 994 TGTATTGGAGACTTAATCGGAGTCCACACCGAGCTTCAGAGTGGAGGATTCATTTGT 1053  
DB 259 TGTATTGGAGACTTAATCGGAGTCCACACCGAGCTTCAGAGTGGAGGATTCATTTGT 200  
QY 1054 ACCCAGAATTGGCAATTTTACCAAGCATTTCAAGGATTAGTATTATCTATCAAGAGCTGT 1113  
DB 199 ACCCAGAATTGGCAATTTTACCAAGCATTTCAAGGATTAGTATTATCTATCAAGAGCTGT 140  
QY 1114 AAGTAACTTGGTGAAGGACACAGGTACTATCATTTGAAAACCTTGACAATGGGTCTTCT 1173

DB 139 AAGTAACTTGGTGAAGGACACAGGTACTATCATTTGAAAACCTTGACAATGGGTCTTCT 80  
QY 1174 TCCATTACACCTCTCTTATATTTTAAAGGCGCTGTGAATATATCTTATTAACCTGCATATCAC 1233  
DB 79 TCCATTACACCTCTCTTATATTTTAAAGGCGCTGTGAATA-ATTATTAACCTGCATATCAC 21  
QY 1234 AAAATAAAACATATTTCTCT 1253  
DB 20 AAAATAAAACATTTTCTCT 1

## RESULT 8

BGI49286/c

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Sequencing by: Greg Lennon, Ph.D.

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to:

info@image.llnl.gov

Seq primer: -40UP from Gibco

High quality sequence stop: 416.

Location/Qualifiers

1. .609

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:3366717"

/tissue\_type="carcinoid"

/lab\_host="DH10B"

/clone\_lib="NCI CGAP Lu24"

/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; plasmid DNA from the normalized library NCI CGAP Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clones 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo."

Query Match 45.8%; Score 580.4; DB 12; Length 609;  
Best Local Similarity 97.4%; Pred. No. 1.2e-151;  
Matches 590; Conservative 0; Mismatches 16; Indels 0; Gaps 0;  
  
QY 650 TCATTTCATGCCCCAGCTGTGCAACGGGCGAGCTCATCAGAGTTCCTGCGAGGCTCC 709  
DB 608 TCATTTCATGCTCCAGCTGTGCAACGGGCGAGCTCATCAGAGTTCCTGCGAGGCTCA 549  
QY 710 TCACCACACTTTCAGTCGGCCCGCCAGGACAAAATTCCTCCATTTTCCAAAGTCGGATCTT 769  
DB 548 TCACCACACTTTCAGTCGGCCCGCCAGGACAAAATTCCTCCATTTTCCAAAGTCGGATCTC 489

## ORIGIN

Query Match 45.8%; Score 580.4; DB 12; Length 609;  
Best Local Similarity 97.4%; Pred. No. 1.2e-151;  
Matches 590; Conservative 0; Mismatches 16; Indels 0; Gaps 0;  
  
QY 650 TCATTTCATGCCCCAGCTGTGCAACGGGCGAGCTCATCAGAGTTCCTGCGAGGCTCC 709  
DB 608 TCATTTCATGCTCCAGCTGTGCAACGGGCGAGCTCATCAGAGTTCCTGCGAGGCTCA 549  
QY 710 TCACCACACTTTCAGTCGGCCCGCCAGGACAAAATTCCTCCATTTTCCAAAGTCGGATCTT 769  
DB 548 TCACCACACTTTCAGTCGGCCCGCCAGGACAAAATTCCTCCATTTTCCAAAGTCGGATCTC 489

```
770 TTCTTGACGACATCTTTGCGAGCTGGATGGCTCAACGGCTGACGACACACCTTGTTACAG 829
Db      |||||
488 TTCTTGATGATCATCTTTGCGAGCTGGATGGATCAACGGCTGACGACACACCTTGTTACAG 429
QY      |||||
830 AAACCTGGCGAGCGAAAGACAGAGCTTCTTCAAACTGCTCCCTTCCCTTACCATGCT 889
Db      |||||
428 AAACCTGGCGAGCGAAAGACAGAGCTTCTTCAAACTGCTCCCTTCCCTTACCATGCT 369
QY      |||||
890 ACATATATAAAGCAATTAATATATACGACACTCTTATTTTCAGTTCCTTATCAAGATCAG 949
Db      |||||
368 ACATATATAAAGCAATTAATATATACGACACTCTTATTTTCAGTTCCTTATCAAGATCAG 309
QY      |||||
950 CCAAGTGGTGTATTTCCCAAGAGCGACCAAAATCGCTGACATGTATTTGGAGACCTAA 1009
Db      |||||
308 CCAAGTGGTGTATTTCCCAAGAGCGACCAAAATCGCTGACATGTATTTGGAGACCTAA 249
QY      |||||
1010 ATCGGAGTCCACACCAAGCCCTTCAGAGTGGAGGATTCATTTGTACCCAGAAATGGCAAA 1069
Db      |||||
248 ATCGGAGTCCACACCAAGCCCTTCAGAGTGGAGGATTCATTTGTACCCAGAAATGGCAAA 189
QY      |||||
1070 TTTACCAAGCATTTCAAGGATAGTATATATATATGTAAGCTGTAAGTAACTTTGGTGAA 1129
Db      |||||
188 TTTACCAAGCATTTCAAGGATAGTATATATATATGTAAGCTGTAAGTAACTTTGGTGAA 129
QY      |||||
1130 AGGACACAGGTACTATCATTTGAAACCTTGGACATGGGTCTTCTTCCATTACACCTTCTT 1189
Db      |||||
128 AGGACACAGGTACTATCATTTGAAACCTTGGACATGGGTCTTCTTCCATTACACCTTCTT 69
QY      |||||
1190 TATATTTTAAAGGCTGTGAATATATATATATATATATATATATATATATATATATATAT 1249
Db      |||||
68 TATATTTTAAAGGCTGTGAATATATATATATATATATATATATATATATATATATATAT 9
QY      |||||
1250 CTCTCA 1255
Db      |||||
8 CTCTCA 3
```

```
RESULT 9
AI420898/c
LOCUS
DEFINITION
  tf03g06.x1 NCI CGAP Pr28 Homo sapiens cDNA clone IMAGE:2095162 3'
  similar to TR:000115 O00115 HYPOTHETICAL HUMAN PROTEIN R31240_2.1;
  mRNA sequence.
ACCESSION
  AI420898
VERSION
  AI420898.1
KEYWORDS
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 554)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
  Emmert-Buck, M.D., Ph.D.
  cDNA Library preparation: M. Bento Soares, Ph.D.
  cDNA Library Arrayed by: Greg Lennon, Ph.D.
  DNA Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/BLNL at:
  www.bio.llnl.gov/bbrp/image/image.html
  Insert Length: 1310 Std Error: 0.00
  Seq primer: -40up from Gibco
  High quality sequence stop: 431.
  Location/Qualifiers
  1..554
  /organism="Homo sapiens"
  /mol_type="mRNA"
```

```
/db xref="taxon:9606"
/clone="IMAGE:2095162"
/sex="male"
/dev stage="adult"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Pr28"
/notes="Organ: prostate; Vector: pT73D-Pac (Pharmacia)
with a modified polylinker; Plasmid DNA from the
normalized library NCI CGAP Pr22 was prepared, and ss
circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneIDs
985608-986759, 1101192-1101959, and 1217928-1220615).
Subtraction by Bento Soares and M. Fatima Bonaldo."
```

```
ORIGIN
Query Match 43.6%; Score 553; DB 9; Length 554;
Best Local Similarity 99.8%; Pred. No. 5.8e-144;
Matches 553; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 715 ACACCTTCAGTCGGCCAGGACACAAAATTCCTCCATTTTGCAAAAGTCGGATTTCTTTCTT 774
Db      |||||
554 ACACCTTCAGTCGGCCAGGACACAAAATTCCTCCATTTTGCAAAAGTCGGATTTCTTTCTT 495
QY 775 GACGACATCTTTTGCGAGCTGGATGGCTCAACGGCTGAAGACACACTTGTTAACGAAACC 834
Db      |||||
494 GACGACATCTTTTGCGAGCTGGATGGCTCAACGGCTGAAGACACACTTGTTAACGAAACC 435
QY 835 TGGCAGCGAAAAGACGACAGCTTCCTTCAAACTGCTCCCTTCCATTACCATGCTTACCAAT 894
Db      |||||
434 TGGCAGCGAAAAGACGACAGCTTCCTTCAAACTGCTCCCTTCCATTACCATGCTTACCAAT 375
QY 895 ATAAAGCAATTAATATATATATATATATATATATATATATATATATATATATATATATAT 954
Db      |||||
374 ATAAAGCAATTAATATATATATATATATATATATATATATATATATATATATATATATAT 315
QY 955 TGGTGTATTTCCCAAAAGGGCACCAAAATCGCTGGACATGTATTTGGAGACCTTAATCGG 1014
Db      |||||
314 TGGTGTATTTCCCAAAAGGGCACCAAAATCGCTGGACATGTATTTGGAGACCTTAATCGG 255
QY 1015 AGTCACACAGAGCTTCAGAGTGGAGGATTCATTTGACCCAGAAATGGCAAAATTTAC 1074
Db      |||||
254 AGTCACACAGAGCTTCAGAGTGGAGGATTCATTTGACCCAGAAATGGCAAAATTTAC 195
QY 1075 CAAGCATTTCAAGGATTAATATATATATATATATATATATATATATATATATATATATAT 1134
Db      |||||
194 CAAGCATTTCAAGGATTAATATATATATATATATATATATATATATATATATATATATAT 135
QY 1135 ACAGTACTATATATATATATATATATATATATATATATATATATATATATATATATATAT 1194
Db      |||||
134 ACAGTACTATATATATATATATATATATATATATATATATATATATATATATATATATAT 75
QY 1195 TTTAAAGGCTGTGAATATATATATATATATATATATATATATATATATATATATATATAT 1254
Db      |||||
74 TTTAAAGGCTGTGAATATATATATATATATATATATATATATATATATATATATATATAT 15
QY 1255 ATGTTTACCATTTA 1268
Db      |||||
14 ATGTTTACCATTTA 1
```

```
RESULT 10
AW295184/c
LOCUS
DEFINITION
  UI-H-B12-aib-e-05-0-UI.sl NCI CGAP_Sub4 Homo sapiens cDNA clone
  IMAGE:2728785 3', mRNA sequence.
ACCESSION
  AW295184
VERSION
  AW295184.1
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 571)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps@mail.nih.gov](mailto:cgaps@mail.nih.gov)  
 The sequence contained an oligo-dT track that was present in the  
 oligonucleotide that was used to prime the synthesis of first  
 strand cDNA and therefore this may represent a bonafide poly A  
 tail. cDNA library Preparation: M.B. Soares Lab Clone Distribution:  
 NCI-CGAP clone distribution information can be found through the  
 I.M.A.G.E. Consortium/ILNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
 Seq primer: M13 Forward  
 POLYA-Yes.

Location/Qualifiers  
 1..571  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2728785"  
 /lab\_host="DRI0B (Life Technologies)"  
 /clone\_lib="NCI CGAP Sub4"  
 /note="vector: p7T3D-Pac (Pharmacia) with a modified  
 polylinker; Site 1: Not I; Site 2: Eco RI; The  
 NCI CGAP Sub4 library is a subtracted library derived from  
 the NCI CGAP Sub2 library which is a subtracted library  
 derived from the NCI CGAP Sub1 library, which is a  
 subtracted library derived from BI. BI constitutes a  
 mixture of 21 normalized or subtracted NCI CGAP  
 libraries: NCI CGAP Co4, NCI CGAP Pr22, NCI CGAP Pr28,  
 NCI CGAP Co10, NCI CGAP Co16, NCI CGAP Kids,  
 NCI CGAP Kid2, NCI CGAP Kid3, NCI CGAP Kid11,  
 NCI CGAP Lym2, NCI CGAP Br2, NCI CGAP Co8, NCI CGAP CLL1,  
 NCI CGAP Lei2, NCI CGAP Brn23, NCI CGAP Lu5,  
 NCI CGAP Lu24, NCI CGAP Lu19, NCI CGAP GC4, NCI CGAP GC6,  
 NCI CGAP Brn25. These 21 libraries were pooled and a  
 single-stranded DNA preparation of the resulting mixture  
 was used as a tracer in a subtractive hybridization with  
 a driver whose composition is detailed below:  
 NCI CGAP Kid3 pool 1: LLAM 3334-3337, 3682-3683,  
 3798-3803 (IMAGE Clonids 1322376-1323911,  
 1456008-1456775, 1500552-1502855) NCI CGAP Kid5 pool 1:  
 LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE Clonids  
 1323912-1325831, 1471368-1472903, 1492104-1493255)  
 NCI CGAP Lu5 pool 1: LLAM 3575-3582, 3851-3854 (IMAGE  
 Clonids 1414920-1417991, 1520904-1522439) NCI CGAP GC4  
 pool 1: LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE  
 Clonids 1257096-1258631, 1469064-1470983,  
 1475592-1476743) NCI CGAP Pr22 pool 1: LLAM 2457-2459,  
 2758-2759, 3062-3068 (IMAGE Clonids 985608-986759,  
 1101392-1101959, 1217928-1220615) NCI CGAP Co10 pool 1:  
 LLAM 2644-2653, 2871-2872 (IMAGE Clonids 1057416-1061255,  
 1144584-1145351) Subtraction was performed as previously  
 described [Bonaldi, Lennon & Soares (1996): Normalization  
 and Subtraction: Two Approaches To Facilitate Gene  
 Discovery. Genome Research 6, 791-806.]  
 TAG\_TISSUE=lung  
 TAG\_LIB=NCI CGAP\_Lu5  
 TAG\_SEQ=CAAC

ORIGIN  
 Query Match 43.4%; Score 550.2; DB 10; Length 571;  
 Best Local Similarity 99.5%; Pred. No. 3.6e-143;  
 Matches 552; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 701 GCAGGCTCTCCACACATTCAGTCGGCCAGGACAAAATTCCTCCATTTGCAAGT 760  
 Db 571 GCAGGCTCTCCACACATTCAGTCGGCCAGGACAAAATTCCTCCATTTGCAAGT 512  
 QY 761 CGGATCTTTTCTTGACGATCTTTGCGCCCTGGATGGCTCAACGGCTGAAGACACACT 820

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

FEATURES  
 source

11  
 AI694455/c  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 536)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps@mail.nih.gov](mailto:cgaps@mail.nih.gov)  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/ILNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
 Insert length: 989 Std Error: 0.00  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 453.  
 Location/Qualifiers  
 1..536  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2338243"  
 /tissue\_type="carcinoid"

Db 511 CGAATCTTTCTTGACATCTTTGAGCCTGGATGGCTCAACGGCTGAAGACACT 452  
 QY 821 TGTAAACAGAAACCTGGCAGCGAAAAGACAAAGAGCTTCCTTCAAACCTGCTCCCTTCCTT 880  
 Db 451 TGTAAACAGAAACCTGGCAGCGAAAAGACAAAGAGCTTCCTTCAAACCTGCTCCCTTCCTT 392  
 QY 881 ACCATGCTCAATATAAAGCAATTAATATATATATATATATATATATATATATATATATAT 940  
 Db 391 ACCATGCTCAATATAAAGCAATTAATATATATATATATATATATATATATATATATATAT 332  
 QY 941 AAGATCAGCCCAAGTGTATTTCCCAAAGGCGACCAAAATTCGCTGGACATGATTG 1000  
 Db 331 AAGATCAGCCCAAGTGTATTTCCCAAAGGCGACCAAAATTCGCTGGACATGATTG 272  
 QY 1001 GAGACCTAAATCGGAGTCCACCAAGCCCTTCAGAACTGGAGGATTCATTTGTATCCACGA 1060  
 Db 271 GAGACCTAAATCGGAGTCCACCAAGCCCTTCAGAACTGGAGGATTCATTTGTATCCACGA 212  
 QY 1061 ATTGGCAATTTACCAAGCAATTTCAAGATTAATATATATATATATATATATATATATATAT 1120  
 Db 211 ATTGGCAATTTACCAAGCAATTTCAAGATTAATATATATATATATATATATATATATATAT 152  
 QY 1121 CTGTGTGAAGGACACAGGTACTATCATTTGAAACCTTGACAACTGGAGGATTCATTTGTATCC 1180  
 Db 151 CTGTGTGAAGGACACAGGTACTATCATTTGAAACCTTGACAACTGGAGGATTCATTTGTATCC 92  
 QY 1181 CACCTCTTTATATTTTAAAGGCTGTGATATATATATATATATATATATATATATATATATAT 1240  
 Db 91 CACCTCTTTATATTTTAAAGGCTGTGATATATATATATATATATATATATATATATATATAT 32  
 QY 1241 AACATATTTCTCTCA 1255  
 Db 31 AACATTTTCTCTCA 17

RESULT 11  
 AI694455/c  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 536)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps@mail.nih.gov](mailto:cgaps@mail.nih.gov)  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/ILNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
 Insert length: 989 Std Error: 0.00  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 453.  
 Location/Qualifiers  
 1..536  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2338243"  
 /tissue\_type="carcinoid"

FEATURES  
 source

/lab host="DH10B"  
/clone\_lib="NCI\_CGAP\_Lu24"  
/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 41.6%; Score 528; DB 9; Length 536;  
Best Local Similarity 99.1%; Pred. No. 6.1e-137;  
Matches 531; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 718 CTTGAGTCGGCCAGGCAAAATTCCTCAATTTGCAAAAGTCGGATCTTTCTTGAC 777  
DB CTTGAGTCGGCCAGGCAAAATTCCTCAATTTGCAAAAGTCGGATCTTTCTTGAT 477

QY 778 GACATCTTTGAGCCTGATGCTCAACGGCTGAAGACACACTTGTTHAAGAACCTGG 837  
DB GACATCTTTGAGCCTGATGCTCAACGGCTGAAGACACACTTGTTHAAGAACCTGG 417

QY 838 CAGCGAAAGACAGAGCTTCCTCAACCTGCTCCCTTCCCTTACCATGTCTACAAATA 897  
DB CAGCGAAAGACAGAGCTTCCTCAACCTGCTCCCTTCCCTTACCATGTCTACAAATA 357

QY 898 AAAGCAATTAATATACGACACTCTTATTTTCAGTCTTATCAAGATCAGCGCAAGTGG 957  
DB AAAGCAATTAATATACGACACTCTTATTTTCAGTCTTATCAAGATCAGCGCAAGTGG 297

QY 958 TGTATTTCCAAAGGGCAACAAATCGCTGGAATGATTTGGAGACCTTAATCGGAGT 1017  
DB TGTATTTCCAAAGGGCAACAAATCGCTGGAATGATTTGGAGACCTTAATCGGAGT 237

QY 1018 CCACCAAGCTTCAGAGTGGAGGATTCATTTGTCACAGATTTGCAAAATTTACCAA 1077  
DB CCACCAAGCTTCAGAGTGGAGGATTCATTTGTCACAGATTTGCAAAATTTACCAA 177

QY 1078 GCATTTCAAGGATTAGTATTATATATGAAAGCTGTAAGTAACTTGTGAAAGACACA 1137  
DB GCATTTCAAGGATTAGTATTATATGAAAGCTGTAAGTAACTTGTGAAAGACACA 117

QY 1138 GGTACTATCATTTGAAAACCTTGCAATGGTCTCTTCCATTTACACCTTCTTTATTTT 1197  
DB GGTACTATCATTTGAAAACCTTGCAATGGTCTCTTCCATTTACACCTTCTTTATTTT 57

QY 1198 AAAGGCTGTGATATATATTAACCTGCTATATCAAAATATAAATATTTCTCT 1253  
DB AAAGGCTGTGATATATATTAACCTGCTATATCAAAATATAAATATTTCTCT 1

RESULT 12  
AI820599/c  
LOCUS  
DEFINITION  
os07c09.x5 NCI CGAP Lu5 Homo sapiens cDNA clone IMAGE1604656 3', similar to TR:090759 090759 DEOXYRIBONUCLEASE II ; mRNA sequence.  
AI820599  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 521)  
REFERENCE  
AUTHORS  
TITLE  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL  
Unpublished (1997)  
Other ESTs: os07c09.y5  
Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LiNL at:  
www-bio.llnl.gov/bbrp/image/image.html

This read is a RESEQUENCE of a previously sequenced human clone  
Original clone citation: National Cancer Institute, Cancer Genome  
Anatomy Project (CGAP), Tumor Gene Index  
This read has been verified (found to hit its original self in the  
correct orientation)  
Insert Length: 947 Std Error: 0.00  
Seq primer: -400P from Gibco  
High quality sequence stop: 460.

FEATURES  
source

1. 521  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1604656"  
/tissue\_type="carcinoid"  
/lab host="DH10B"  
/clone\_lib="NCI CGAP Lu5"  
/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared with a neuroendocrine lung carcinoid, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 40.3%; Score 511.4; DB 9; Length 521;  
Best Local Similarity 98.8%; Pred. No. 2.8e-132;  
Matches 515; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 733 GGACAAATTTCTCCATTTTGCAGATCGGATCTTTCTTTCGACGACATCTTTCAGCC 792  
DB GGGACAAATTTCTCCATTTTGCAGATCGGATCTTTCTTTCGACGACATCTTTCAGCC 462

QY 793 TGGATGGCTCAACGGCTCAAGACACACATTTGTTAAACAGAAACCTGGCAGCAAAAGACAA 852  
DB TGGATGGCTCAACGGCTCAAGACACACATTTGTTAAACAGAAACCTGGCAGCAAAAGACAA 402

QY 853 GAGCTTCTCTCAAACCTGCTCCCTTCCATGCTCAATATATAAAGCAATTAATTA 912  
DB GAGCTTCTCTCAAACCTGCTCCCTTCCATGCTCAATATATAAAGCAATTAATTA 342

QY 913 TCACGACACTCTTATTTTCAAGTTCATCAAGATCAAGCTGTTGTTATTTTCCCAAAAG 972  
DB TCACGACACTCTTATTTTCAAGTTCATCAAGATCAAGCTGTTGTTATTTTCCCAAAAG 282

QY 973 GGCACCAAAATTCGCTGGACATGATTTGGAGACCTTAATTCGGAGTCCACCAAGCTTC 1032  
DB GGCACCAAAATTCGCTGGACATGATTTGGAGACCTTAATTCGGAGTCCACCAAGCTTC 222

QY 1033 AGAAGTGGAGGATTCATTTGTTACCGAATTCGGCAATTTACAGCATTTCAAGGATTA 1092  
DB AGAAGTGGAGGATTCATTTGTTACCGAATTCGGCAATTTACAGCATTTCAAGGATTA 162

QY 1093 GTATTATATCTAAGAGCTGTAAGTAACTTCGTGAAAGACACACAGGTACTATCATTTGAA 1152  
DB GTATTATATCTAAGAGCTGTAAGTAACTTCGTGAAAGACACACAGGTACTATCATTTGAA 102

QY 1153 AACCTTGAACATGGGTCTTCTTCCATTCACACCTTCTTTATATTTTAAAGCCTGTGAATA 1212  
DB AACCTTGAACATGGGTCTTCTTCCATTCACACCTTCTTTATATTTTAAAGCCTGTGAATA 42





Library NCI CGAP Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo.

ORIGIN		Query Match	36.9%;	Score 467.8;	DB 13;	Length 482;	
		Best Local Similarity	99.4%;	Pred. No. 4.9e-120;			
		Matches 480;	Conservative	0;	Mismatches 2;	Indels 1;	Gaps 1;
Qy	335	AGAGTAAACACACAGCCTATCTAATATACATGATGAGTCCCTAAACCTGTGAATTACA	394				
Db	1	AGAGTAAACACACAGCCTATCTAATATACATGATGAGTCCCTAAACCTGTGAATTACA	60				
Qy	395	GTAGAAAGTATGGACACACAAAGTTTACTGTGTGGAACAGAGTTCAAGGGTTCGGC	454				
Db	61	GCAGAAAGTATGGACACACAAAGTTTACTGTGTGGAACAGAGTTCAAGGGTTCGGC	120				
Qy	455	TGATTTCATTCCTCCCTCAGTTTCCTCAATTCGGGAAGAGGCTATGATTATCCACCCA	514				
Db	121	TGATTTCATTCCTCCCTCAGTTTCCTCAATTCGGGAAGAGGCTATGATTATCCACCCA	180				
Qy	515	CAGGAGACGAAATGGACAAAGTGCATCTGCATTAACCTTCAAGTACCAACAGTATGAG	574				
Db	181	CAGGAGACGAAATGGACAAAGTGCATCTGCATTAACCTTCAAGTACCAACAGTATGAG	240				
Qy	575	CAATAGATTCTCAGCTCTTGCTGCAACCCCAACCTCTATAGTGTCTCCATCCAGCCA	634				
Db	241	CAATAGATTCTCAGCTCTTGCTGCAACCCCAACCTCTATAGTGTCTCCATCCAGCCA	300				
Qy	635	CTTTTACCAGAGCTCAATTCATGCCCCAGCTGTGCACAGGCCAGCTCATCAGAGA	694				
Db	301	CTTTTACCAGAGCTCAATTCATGCCCCAGCTGTGCACAGGCCAGCTCATCAGAGA	360				
Qy	695	TTCTGGGAGCTCTCTACACACACTTCAGTGGCCCCAGGACAAATTCCTCCATTTTG	754				
Db	361	TTCTGGGAGCTCTCTACACACACTTCAGTGGCCCCAGGACAAATTCCTCCATTTTG	419				
Qy	755	CAAGTCGGATTCCTTTCTTGACGACATCTTTGACAGCTGTGATGCTCAACGGCTGAAGA	814				
Db	420	CAAGTCGGATTCCTTTCTTGACGACATCTTTGACAGCTGTGATGCTCAACGGCTGAAGA	479				
Qy	815	CAC 817					
Db	480	CAC 482					

RESULT 15  
AA988125/c  
LOCUS  
DEFINITION  
os07c09.s1 NCI CGAP Lu5 Homo sapiens CDNA clone IMAGE:1604656 3', similar to TR:000115 000115 HYPOTHETICAL HUMAN PROTEIN R31240\_2. ; mRNA sequence.  
AA988125  
AA988125.1 GI:3173489  
EST.  
Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 446)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapps-remail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/hbrp/image/image.html  
Insert length: 947 Std Error: 0.00  
Seq primer: -40m13 fwd. ET from Amersham  
High quality sequence stop: 365.

FEATURES		Location/Qualifiers
source		1..446
		/organism="Homo sapiens"
		/mol_type="mRNA"
		/db_xref="taxon:9606"
		/clone="IMAGE:1604656"
		/tissue_type="carcinoid"
		/lab_host="DH10B"
		/clone_lib="NCI CGAP Lu5"
		/note="Organ: lung; Vector: pTV73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was then primed with a neuroendocrine lung carcinoid, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTV73 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN		Query Match	34.8%;	Score 441.8;	DB 9;	Length 446;	
		Best Local Similarity	99.3%;	Pred. No. 9.6e-113;			
		Matches 443;	Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;
Qy	808	CTGAAGACACACTTGTTAACAGAAACCTGGCAGCGAAAGACAGAGCTTCTTCAAAC	867				
Db	446	CTGAAGACACACTTGTTAACAGAAACCTGGCAGCGAAAGACAGAGCTTCTTCAAAC	387				
Qy	868	TGCTCCCTTCTTACCATGCTCAATATAAAGCAATTAATATCAGCACCTCTTAT	927				
Db	386	TGCTCCCTTCTTACCATGCTCAATATAAAGCAATTAATATCAGCACCTCTTAT	327				
Qy	928	TTCAGTCTTATCAAGATCAGCCAGTGGTGTATTTCCCAAGGGCACCAGAAATGCG	987				
Db	326	TTCAGTCTTATCAAGATCAGCCAGTGGTGTATTTCCCAAGGGCACCAGAAATGCG	267				
Qy	988	TGGACATGATTTGGAGACCTAAATCGGAGTCCACACAGCCTTCAGAAAGTGGAGATTC	1047				
Db	266	TGGACATGATTTGGAGACCTAAATCGGAGTCCACACAGCCTTCAGAAAGTGGAGATTC	207				
Qy	1048	ATTGTACCCAGAAATGGCAATTTTACCAAGCATTTCAAGGATTTAGTATTACTATGAA	1107				
Db	206	ATTGTACCCAGAAATGGCAATTTTACCAAGCATTTCAAGGATTTAGTATTACTATGAA	147				
Qy	1108	AGCTGTAAGTAACTTGGTGAAGGACACAGGTACTATCATTTGAAAACCTTGACAATGG	1167				
Db	146	AGCTGTAAGTAACTTGGTGAAGGACACAGGTACTATCATTTGAAAACCTTGACAATGG	87				
Qy	1168	TCTTCTTCCATFACACCTTCTTATTTTAAAGCCGTGTGAATATATCTTATAACCTGCA	1227				
Db	86	TCTTCTTCCATFACACCTTCTTATTTTAAAGCCGTGTGAATATATCTTATAACCTGCA	27				
Qy	1228	TATCAAAAATAAATATATTTCTCT	1253				
Db	26	TATCAAAAATAAATATATTTCTCT	1				

Search completed: October 14, 2004, 05:24:16  
Job time : 3099.15 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 22:11:29 ; Search time 71.8017 Seconds  
(without alignments)  
1404.834 Million cell updates/sec

Title: US-10-790-589-4  
Perfect score: 1933  
Sequence: 1 MMARLRTSFALLFLGLFGV.....QNWQIQAFQGLVLYESCK 357

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1933	100.0	357	4	AAE13015 Human deo
2	1933	100.0	357	5	ABG61821 Prostate
3	1926	99.6	361	4	AAE13015 Human deo
4	1293.5	66.9	354	4	AAE13015 Human deo
5	1293.5	66.9	354	4	AAE13015 Human deo
6	613	31.7	360	2	AAW31495 Human DNA
7	613	31.7	360	2	AAW31495 Human DNA
8	459	23.7	276	2	AAW37921 Bos tauru
9	449	23.2	366	4	ABG63496 Drosophil
10	229	11.8	192	6	ABG72082 Mouse end
11	98.5	5.1	391	6	ABU41838 Protein e
12	98.5	5.1	702	2	ABU41838 Protein e
13	98.5	5.1	1272	4	ABG03108 Chlamydia
14	97.5	5.0	697	5	ABG91036 Chlamydia
15	97.5	5.0	697	7	ADDA3730 Chlamydia
16	97	5.0	426	2	AAE13015 Human deo
17	96.5	5.0	397	4	ABG11583 Raphanus
18	96.5	5.0	661	2	ABG11583 Raphanus
19	95.5	4.9	308	5	ABP29969 Streptoco
20	95.5	4.9	328	5	ABP27445 Streptoco
21	95	4.9	595	2	AAE13015 Human deo
22	95	4.9	595	2	AAW54157
23	95	4.9	595	2	AAW87487
24	95	4.9	681	5	ABW77755
25	95	4.9	687	5	AAW47670 MOL9b fra

26	95	4.9	704	4	AAE13015
27	95	4.9	704	5	AAE13015
28	95	4.9	704	5	AAE13015
29	95	4.9	704	7	AAE13015
30	95	4.9	714	5	AAU74744
31	95	4.9	970	6	ABR53403
32	94.5	4.9	2295	3	AAE18180
33	93	4.8	526	6	ABU48992
34	93	4.8	2472	5	AAU98502
35	93	4.8	6576	6	ABO14699
36	93	4.8	6669	6	ABO14698
37	93	4.8	6669	7	ADD46052
38	93	4.8	6700	6	ABO14697
39	92.5	4.8	638	5	AAW47669
40	92	4.8	517	4	ABB66825
41	92	4.8	524	4	ABB69124
42	92	4.8	641	7	ADB99981
43	92	4.8	641	7	ADC00607
44	91	4.7	704	5	ABP69720
45	91	4.7	1035	4	AAW39241

## ALIGNMENTS

RESULT 1  
AAE13015

ID AAE13015 standard; protein; 357 AA.

XX AC AAE13015;

XX AC 28-JAN-2002 (first entry)

XX DE Human deoxyribonuclease (DNase) II beta protein.

XX KW Human; deoxyribonuclease; DNase II beta protein; mucous plug; lung;  
cystic fibrosis; pulmonary; chromosome 1p22.

XX OS Homo sapiens.

XX PN WO200175082-A1.

XX PD 11-OCT-2001.

XX PF 02-APR-2001; 2001WO-US010635.

XX PR 03-APR-2000; 2000US-00541840.

XX PR 19-MAY-2000; 2000US-00574942.

XX PR (DART-) DARTMOUTH COLLEGE.

XX PR Eastman AR, Krieser RJ;

XX PR WPI; 2001-662972/76.

XX PR N-PSDB; AAD21289.

XX PT New cDNA encoding a deoxyribonuclease II beta enzyme useful for degrading  
DNA present in the mucous plugs in the lungs of cystic fibrosis patients.

XX PS Claim 5; Page 16-17; 21pp; English.

XX CC The invention relates to deoxyribonuclease (DNase) II beta proteins and  
their corresponding cDNAs. The DNase II beta may be useful to digest DNA  
in the mucous plugs in lungs of cystic fibrosis patients and so reduce  
their viscosity. The present sequence is human DNase II beta protein. The  
human DNase II beta gene is located at chromosome 1p22

XX SQ Sequence 357 AA;

Query Match

Best Local Similarity 100.0%; Score 1933; DB 4; Length 357;

Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMARLLRTSFALLFLGLFGVLGAATISCRNEEGKAVDWFYKLPKQNKESGETGLEYL 60  
 DB 1 MMARLLRTSFALLFLGLFGVLGAATISCRNEEGKAVDWFYKLPKQNKESGETGLEYL 60  
 QY 61 YLDSTTSWRKSEQLMNDTKSVLGRITLQQLYEAVASKSNNTAYLIYNDGVKPKVNSRKY 120  
 DB 61 YLDSTTSWRKSEQLMNDTKSVLGRITLQQLYEAVASKSNNTAYLIYNDGVKPKVNSRKY 120  
 QY 121 GHTKGLLLMNRVQGFLLHSLIPQPPPIPEEGYDYPPTGRRNGQSGICITFKKNOYEIDS 180  
 DB 121 GHTKGLLLMNRVQGFLLHSLIPQPPPIPEEGYDYPPTGRRNGQSGICITFKKNOYEIDS 180  
 QY 181 QLLVCNPNVYSCSIPATPHQELIHPOLCTRASSEIPGRLLITLQSAQOGKFLHFAKSD 240  
 DB 181 QLLVCNPNVYSCSIPATPHQELIHPOLCTRASSEIPGRLLITLQSAQOGKFLHFAKSD 240  
 QY 241 SFLLDDIFAAMWAQRKTHLLTETWQRKQELPNSCLPYHYVNIKAIKLSRHSYFSSYQD 300  
 DB 241 SFLLDDIFAAMWAQRKTHLLTETWQRKQELPNSCLPYHYVNIKAIKLSRHSYFSSYQD 300  
 QY 301 HAKWCISQKTKNRWTCIGDLNRSPOHAFRSGGFICTQNWQIYQAFQGLVLYYESCK 357  
 DB 301 HAKWCISQKTKNRWTCIGDLNRSPOHAFRSGGFICTQNWQIYQAFQGLVLYYESCK 357

RESULT 2

ABG61821  
 ID ABG61821 standard; protein; 357 AA.

AC ABG61821;

DT 15-AUG-2002 (first entry)

DE Prostate cancer-associated protein #22.

KE Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.

OS Mammalia.

PN WO200230269-A2.

PD 18-APR-2002.

XX 12-OCT-2001; 2001WO-US032045.

XX 13-OCT-2000; 2000US-00687576.

XX 08-DEC-2000; 2000US-00733288.

XX 08-DEC-2000; 2000US-00733742.

XX 24-JAN-2001; 2001US-0263957P.

XX 16-MAR-2001; 2001US-0276791P.

XX 16-MAR-2001; 2001US-0276888P.

XX 06-APR-2001; 2001US-0281922P.

XX 24-APR-2001; 2001US-0286214P.

XX 30-APR-2001; 2001US-00847046.

XX 04-MAY-2001; 2001US-0288589P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;

XX WPI; 2002-471335/50.

XX N-PSDB; ABK92136.

XX Detecting a prostate cancer-associated transcript in a cell in a patient,  
 XX useful for diagnosing prostate cancer (PC) or screening modulators of PC,  
 XX by determining if prostate cancer-associated genes are expressed in a  
 XX prostate tissue.  
 XX Claim 27; Page 317; 436pp; English.  
 XX The present invention relates to methods of detecting a prostate cancer-  
 XX associated transcript in a cell from a patient. The method comprises  
 XX contacting a biological sample from the patient with prostate cancer-

CC associated polynucleotides (designated PC genes) that selectively  
 CC hybridise to a sequence that is at least 80% identical to them. The  
 CC prostate cancer-associated polynucleotide sequences are differentially  
 CC expressed in prostate tumour tissue or in prostate cancer and are derived  
 CC from the tissues of various organisms such as humans or other mammals  
 CC (e.g. mice, sheep and dogs). The methods of the invention are useful for  
 CC diagnosing and treating prostate cancer in mammals. The prostate cancer-  
 CC associated genes are useful for diagnosing or treating prostate cancer,  
 CC as well as for identifying modulators of prostate cancer or agents that  
 CC inhibit prostate cancer. The nucleic acid sequences are particularly  
 CC useful in gene therapy, as a vaccine or in antisense applications.  
 CC ABG61800-ABG61944 represent prostate cancer-associated proteins  
 XX  
 SQ Sequence 357 AA;

Query Match 100.0%; Score 1933; DB 5; Length 357;

Best Local Similarity 100.0%; Pred. No. 9.1e-174; Indels 0; Gaps 0;  
 Matches 357; Conservative 0; Mismatches 0;

QY 1 MMARLLRTSFALLFLGLFGVLGAATISCRNEEGKAVDWFYKLPKQNKESGETGLEYL 60

DB 1 MMARLLRTSFALLFLGLFGVLGAATISCRNEEGKAVDWFYKLPKQNKESGETGLEYL 60

QY 61 YLDSTTSWRKSEQLMNDTKSVLGRITLQQLYEAVASKSNNTAYLIYNDGVKPKVNSRKY 120

DB 61 YLDSTTSWRKSEQLMNDTKSVLGRITLQQLYEAVASKSNNTAYLIYNDGVKPKVNSRKY 120

QY 121 GHTKGLLLMNRVQGFLLHSLIPQPPPIPEEGYDYPPTGRRNGQSGICITFKKNOYEIDS 180

DB 121 GHTKGLLLMNRVQGFLLHSLIPQPPPIPEEGYDYPPTGRRNGQSGICITFKKNOYEIDS 180

QY 181 QLLVCNPNVYSCSIPATPHQELIHPOLCTRASSEIPGRLLITLQSAQOGKFLHFAKSD 240

DB 181 QLLVCNPNVYSCSIPATPHQELIHPOLCTRASSEIPGRLLITLQSAQOGKFLHFAKSD 240

QY 241 SFLLDDIFAAMWAQRKTHLLTETWQRKQELPNSCLPYHYVNIKAIKLSRHSYFSSYQD 300

DB 241 SFLLDDIFAAMWAQRKTHLLTETWQRKQELPNSCLPYHYVNIKAIKLSRHSYFSSYQD 300

QY 301 HAKWCISQKTKNRWTCIGDLNRSPOHAFRSGGFICTQNWQIYQAFQGLVLYYESCK 357

DB 301 HAKWCISQKTKNRWTCIGDLNRSPOHAFRSGGFICTQNWQIYQAFQGLVLYYESCK 357

RESULT 3

ABG72417  
 ID AAB72417 standard; protein; 361 AA.

AC AAB72417;

XX 03-MAY-2001 (first entry)

DT Human DNase.

DE Human; DNase; DLAD; deoxyribonuclease; endonuclease; cystic fibrosis;  
 XX infectious disease.

XX Homo sapiens.

XX WO200112793-A1.

XX 22-FEB-2001.

XX 01-MAY-2000; 2000WO-JP002893.

XX 17-AUG-1999; 99JP-00230870.

XX (TANU/) TANUMA S.

XX Tanuma S, Shikawa D;

XX WPI; 2001-218348/22.

XX N-PSDB; AAF60716.

XX Acidic deoxyribonuclease capable of divalent cation-independent cleavage  
 PT of DNA under acidic even neutral pH and not inhibited by G-actin, useful  
 PT in remedies for cystic fibrosis and for treatment of infectious diseases.

XX Claim 8; Page 52-53; 61pp; Japanese.

XX The present sequence is a human deoxyribonuclease (DLAD), which is an  
 CC endonuclease. DLAD is capable of divalent cation-independent cleavage of  
 CC DNA under acidic conditions. This protein can be used as a substitute for  
 CC DNase I in treating cystic fibrosis, and is useful in the prevention and  
 CC treatment of infectious diseases

XX SQ Sequence 361 AA;

Query Match 99.6%; Score 1926; DB 4; Length 361;  
 Best Local Similarity 99.7%; Pred. No. 4.2e-173;  
 Matches 356; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MMARLLTSPALLFLGLFGVLGAATISCRNEEGKAVDWFYFKLPKQNKESGETGLEYL 60  
 Db 5 MMARLLTSPALLFLGLFGVLGAATISCRNEEGKAVDWFYFKLPKQNKESGETGLEYL 64  
 QY 61 YLSDTTSRWSKSEQLMNDTKSVLGRITLQOLYEAVASKNNNTAYLIYNDGVPKPNYSRKY 120  
 Db 65 YLSDTTSRWSKSEQLMNDTKSVLGRITLQOLYEAVASKNNNTAYLIYNDGVPKPNYSRKY 124  
 QY 121 GHTKGLLLNVRVQGFWLIHSIPPPPIPERGYDPPPTGRNGOSGICITTKYQYEAIDS 180  
 Db 125 GHTKGLLLNVRVQGFWLIHSIPPPPIPERGYDPPPTGRNGOSGICITTKYQYEAIDS 184  
 QY 181 QLLVCPNVYSCSIPATFHQELHMPQLCTRASSSEIPGRLLTTLQSAQOKFLHFAKSD 240  
 Db 185 QLLVCPNVYSCSIPATFHQELHMPQLCTRASSSEIPGRLLTTLQSAQOKFLHFAKSD 244  
 QY 241 SFLLDIFAAWMAORLKLTHLTETWQRKQELPNSCLPYHYNIKATKLSRHSYFSSYQD 300  
 Db 245 SFLLDIFAAWMAORLKLTHLTETWQRKQELPNSCLPYHYNIKATKLSRHSYFSSYQD 304  
 QY 301 HAKWCISQKTKNRWTCIGDLNRSPOAFRSGGFICTQNWQIYQAFQGLVLYYESCK 357  
 Db 305 HAKWCISQKTKNRWTCIGDLNRSPOAFRSGGFICTQNWQIYQAFQGLVLYYESCK 361

RESULT 4  
 AAB72416  
 ID AAB72416 standard; protein; 354 AA.

AC AAB72416;

XX 03-MAY-2001 (first entry)

DE Murine DNase.

XX Murine; DNase; DLAD; deoxyribonuclease; endonuclease; cystic fibrosis;  
 KW infectious disease.

XX Mus musculus.

XX WO200112793-A1.

XX 22-FEB-2001.

XX 01-MAY-2000; 2000WO-JP002893.

XX 17-AUG-1999; 99JP-00230870.

XX (TANU/) TANUMA S.

XX Tanuma S, Shiohawa D;

XX WPI; 2001-218348/22.

XX N-PSDB; AAF60715.

XX

PT Acidic deoxyribonuclease capable of divalent cation-independent cleavage  
 PT of DNA under acidic even neutral pH and not inhibited by G-actin, useful  
 PT in remedies for cystic fibrosis and for treatment of infectious diseases.

XX Claim 3; Page 48-49; 61pp; Japanese.

XX The present sequence is a murine deoxyribonuclease (DLAD), which is an  
 CC endonuclease. DLAD is capable of divalent cation-independent cleavage of  
 CC DNA under acidic conditions. This protein can be used as a substitute for  
 CC DNase I in treating cystic fibrosis, and is useful in the prevention and  
 CC treatment of infectious diseases

XX SQ Sequence 354 AA;

Query Match 66.9%; Score 1293.5; DB 4; Length 354;  
 Best Local Similarity 66.4%; Pred. No. 2.4e-113;  
 Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

QY 1 MMARLLTSPALLFLGLFGVLGAATISCRNEEGKAVDWFYFKLPKQNKESGETGLEYL 60  
 Db 1 MTAKPLRTVLSLLFFALSGVLGTPEISCRNEYGEAVDWFYFKLPKRTSKASEAGLYL 60  
 QY 61 YLSDTTSRWSKSEQLMNDTKSVLGRITLQOLYEAVASKNNNTAYLIYNDGVPKPNYSRKY 120  
 Db 61 YLSDTTSRWSKSEQLMNDTKSVLGRITLQOLYEAVASKNNNTAYLIYNDGVPKPNYSRKY 119  
 QY 121 GHTKGLLLNVRVQGFWLIHSIPPPPIPERGYDPPPTGRNGOSGICITTKYQYEAIDS 180  
 Db 120 GHAKGLLVNRTQGFWLIHSVPKPPV--HGXEYPTSGRRYVGIGCITTFGYSQFEIDF 177  
 QY 181 QLLVCPNVYSCSIPATFHQELHMPQLCTRASSSEIPGRLLTTLQSAQOKFLHFAKSD 240  
 Db 178 QLLVCPNVYSCSIPATFHQELHMPQLCTRASSSEIPGRLLTTLQSAQOKFLHFAKSD 237  
 QY 241 SFLLDIFAAWMAORLKLTHLTETWQRKQELPNSCLPYHYNIKATKLSRHSYFSSYQD 300  
 Db 238 FYTDDIFTGIAQKLLTHLAAQTWQKKQELPNSCLPYHYNIKLSIGVTSKTSFSSRQD 297  
 QY 301 HAKWCISQKTKNRWTCIGDLNRSPOAFRSGGFICTQNWQIYQAFQGLVLYYESCK 357  
 Db 298 HSKWCVSIGKANRWTCIGDLNRSLOALRGGGFICTQNHYYIYQAFHKLRYRGFCK 354

RESULT 5

AAB13014  
 ID AAB13014 standard; protein; 354 AA.

XX AC AAB13014;

XX 28-JAN-2002 (first entry)

XX Mouse deoxyribonuclease (DNase) II beta protein.

XX Mouse; deoxyribonuclease; DNase II beta protein; mucous plug; lung;  
 KW cystic fibrosis; pulmonary.

XX Mus sp.

XX WO200175082-A1.

XX 11-OCT-2001.

XX 02-APR-2001; 2001WO-US010635.

XX 03-APR-2000; 2000US-00541840.

XX 19-MAY-2000; 2000US-00574942.

XX (DART-) DARTMOUTH COLLEGE.

XX Eastman AR, Krieser RJ;

XX WPI; 2001-662972/76.

DR N-PSDB; AAD21288.  
XX New cDNA encoding a deoxyribonuclease II beta enzyme useful for degrading  
PT DNA present in the mucous plugs in the lungs of cystic fibrosis patients.  
XX  
XX Claim 5; Page 14-15; 21pp; English.  
XX  
XX The invention relates to deoxyribonuclease (DNase) II beta proteins and  
CC their corresponding cDNAs. The DNase II beta may be useful to digest DNA  
CC in the mucous plugs in lungs of cystic fibrosis patients and so reduce  
CC their viscosity. The present sequence is mouse DNase II beta protein  
XX  
XX Sequence 354 AA;  
XX  
Query Match 66.9%; Score 1293.5; DB 4; Length 354;  
Best Local Similarity 66.4%; Pred. No. 2.4e-113;  
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;  
QY 1 MVARLLSTFALLFLGLGVLAGAATISCRNEEGKAVDMFTFYKLPKRONKESGETGLLEYL 60  
Db 1 MTAKPLKRVLSLFFALSGVLGTPEISCRNRYGEADVFIYKLPKRTSKASEAGLQYL 60  
QY 61 YLDSTTRSWRSEQLMNDTKSVLGRTLQOLYEAYASKSNNTAYLIYNDGVPKPVNYSKY 120  
Db 61 YLDSTRTWNKSLYLINSTRSALGRTLQHLYDTHNS-TNDTAYLIYNDGVPGSVNYSRQY 119  
QY 121 GHTGKLLWNRVQGWFLHISIPPEPIPEGYDYPPTGRRNGSGICITFKYNOYEADS 180  
Db 120 GHAKGLLWNRVQGWFLHISIPPEPIPEGYDYPPTGRRNGSGICITFKYNOYEADS 177  
QY 181 QLLVCPNPNVSCSIPATHFOELIHPQLCTRASSSEIPGRLLTTLQSAQGGKFLHFAKSD 240  
Db 178 QLLVCPNPNVSCSIPATHFOELIHPQLCTRASSSEIPGRLLTTLQSAQGGKFLHFAKSD 237  
QY 241 SFLLDDIFAAWMAQRKTHLLTETWQRKQELPNSCLPYHYVNTKAIKLSRHSTVPSYQD 300  
Db 238 FYTDDIFGWIQAQKLTLLAQTQKQKQELPNSCLPYHYVNTKAIKLSRHSTVPSYQD 297  
QY 301 HAKWCISOKGTNRWTCIGDLNRSFHOAFRSGGFTCTQWQIYQAFQGLVLYYESCK 357  
Db 298 HSKWCISOKGTNRWTCIGDLNRSFHOAFRSGGFTCTQWQIYQAFQGLVLYYESCK 354

RESULT 6  
AAW31495  
ID AAW31495 standard; protein; 360 AA.  
XX  
XX AAW31495;  
XX  
XX 11-MAY-1998 (first entry)  
XX  
XX Human DNase II protein.  
DE  
XX DNase II; human; deoxyribonuclease; pulmonary disease; diagnostic;  
KW cellular DNA debris.  
XX  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FH 1. .16  
FT Peptide /label= signal\_peptide  
FT Protein 17. .360  
FT /note= "DNase II"  
FT Modified-site 86. .88  
FT /note= "Asn is N-glycosylated"  
FT Modified-site 212. .214  
FT /note= "Asn is N-glycosylated"  
FT Modified-site 266. .268  
FT /note= "Asn is N-glycosylated"  
FT Modified-site 290. .292  
FT /note= "Asn is N-glycosylated"  
XX  
XX WO9740134-A2.  
PN

XX 30-OCT-1997.  
XX  
XX 23-APR-1997; 97WO-US006664.  
XX  
XX 25-APR-1996; 96US-00639294.  
XX  
XX (GETH ) GENENTECH INC.  
XX  
XX Baker KP, Baron WF;  
XX  
XX WPI; 1997-535820/49.  
XX  
XX N-PSDB; AAV02903.  
XX  
XX Human DNase II and related nucleic acids - useful in protein production  
PT e.g. for therapeutic use to treat systemic lupus erythematosus and  
PT pulmonary diseases e.g. cystic fibrosis.  
XX  
XX Claim 1; Fig 1; 30pp; English.  
XX  
XX This sequence represents a novel human deoxyribonuclease, DNase II. This  
CC protein is useful to reduce the viscoelasticity of DNA-containing  
CC material, e.g. mucus, and used to treat patients with pulmonary diseases  
CC or disorders e.g. cystic fibrosis, or with systemic lupus erythematosus.  
CC It may also improve antibiotic efficacy in the treatment of abscesses,  
CC infected lesions etc., provide treatment in non-infected conditions in  
CC which there is an accumulation of cellular DNA debris, e.g.  
CC pyelonephritis and be used to degrade DNA in biological samples or in  
CC diagnostic assays. The nucleic acids and vectors can be produced  
CC or ex vivo gene therapy, and antisense oligonucleotides can be produced  
CC from the nucleic acids which can bind to and prevent expression of  
CC nucleic acid within cells. The nucleic acids (for a portion) can also be  
CC used for hybridisation assays for nucleic acids encoding human DNase II  
CC in a sample, or to identify and isolate nucleic acids sharing substantial  
CC sequence identity (e.g. encoding naturally-occurring allelic variants of  
CC human DNase II). The antibodies can be used to detect and measure human  
CC DNase II in tissues or clinical samples, and in the purification of human  
CC DNase II. The nucleic acids enable production of human DNase II by  
CC recombinant DNA methods in quantities sufficient for clinical use, not  
CC previously possible  
XX  
XX Sequence 360 AA;  
XX  
Query Match 31.7%; Score 613; DB 2; Length 360;  
Best Local Similarity 36.9%; Pred. No. 5e-49;  
Matches 129; Conservative 61; Mismatches 148; Indels 12; Gaps 6;  
QY 13 LFLGLFGLVGAATISCRNEEGKAVDMFTFYKLPK-RONKESGETGLVYLDSTTRSWRK 71  
Db 4 LLLAALLCVPAGALTCTGDSGQPVDMFVVYKLPALRGSGEAAQRLQYKILDESSGWRD 63  
QY 72 SEQLMNDTKSVLGRTLQOLYEAYASKSNNTAYLIYNDGVPKP-VNYSRKYGHTKGLLW 129  
Db 64 GRALINSEGAVERSLQFLYR---SNTSLAFLIYNDQPPQPSKADSSMRGHTGVLL 120  
QY 130 NRVOGFWLIHSIPQF-PPPIPEGYDYPPTGRRNGSGICITFKYNOYEADQLIVCNPN 188  
Db 121 DHGGFWLHVSVENFPFPAASAAVSWPHSACTYGTQTLTLCVSPFPAQFSGMKGQLTYTPW 180  
QY 189 VYSCSIPATHFOELIHPQLCTRASSSEIPGRLLTTLQSAQGGKFLHFAKSDIFPA 248  
Db 181 VYNYQLGIFAQEFPPDENVKGHVSOEPWNSSITLTSQAGAVFQSFQKFKFGDGLYS 240  
QY 249 AWAQRLKTHLLTETWQRKQELPNSCLPYHYVNTKAIKLSRH-S--YFSSYQDHAKWC 306  
Db 241 GHLAALGTLNLOVQFWHKTVGLTPNSCDIMQVLNVNLIAPPAGPAGPSFNSTEDSHKWC 300  
QY 307 SOKGTNRWTCIGDLNRSFHOAFRSGGFTCTQWQIYQAFQGLVLYYESCK 356  
Db 301 SPKGP---WTCVGMNRNQGEORGGGTILCAQLPALWKAQPLVKNYQPC 347

RESULT 7

	Query Match	31.7%	Score 613	DB 2	Length 360
Best Local Similarity	36.9%	Pred. No. 5e-49			
Matches 129	Conservative	61	Mismatches 148	Indels 12	Gaps 6
QY	13	LFLGLFVLCAANTISCRNERSKAVDWTFYKLPK-RONKSGSETGLEVLYLIDSTRWRK	71		
		:::          :::          :::          :::          :::          :::          :::			
Db	4	LLLAALLCVFAGALITCGDGGOPVDFWFWFKYLPALRGSGEAAQGLQYKYIDEGSGWRD	63		

CC agents, e.g. anticancer agents to promote apoptosis in cells. The  
CC antibodies can also be used to identify cells susceptible to premature  
CC death. The cDNA encoding it is useful to identify agents modulating  
CC apoptosis in cells, by treating cells with an agent, transfecting cells  
CC with cDNA and monitoring apoptosis compared with untreated cells.  
CC Inhibitors identified may be useful in preventing diseases relating to  
CC enhanced chromosomal re-arrangement. Vectors comprising the cDNA can be  
CC used to induce apoptosis in selected cells, e.g. tumour cells or cells  
CC involved in autoimmune disorders. Antisense oligonucleotides can be  
CC administered to cells to inhibit DNase II expression to reduce chromosome  
CC instability associated with cancer  
XX  
SQ Sequence 276 AA;

Query Match 23.7%; Score 459; DB 2; Length 276;  
Best Local Similarity 35.8%; Pred. NO. 1.2e-34;  
Matches 95; Conservative 52; Mismatches 108; Indels 10; Gaps 5;  
QY 98 SNNATYLYNDGVPKPNVY--SRKYGHTKGLLLNVRVQGLIHSIPOPPP-IPEEGYDY 154  
Db 1 NSQAFVLYNDQPKSSSKSSRGHTKGVLLLDQEGFWLIHSVFNPPPRASSAAYSW 60  
QY 155 PPTGRRNGSGGICITFKYNOYEADISQLLCNPNVYSCSIPATHQELIHPQLCTRASS 214  
Db 61 PPGAKYQGLTLCVSPFLTQFLDLSKQLTYTPLYVDRHLEGDPQKPPYLEEVVKGHHV 120  
QY 215 SEIPGRLITLQSAQGGKFLHFAKSDSFLDDIFAAWMAQRKTHLLTETWORKQELPSN 274  
Db 121 RQGPWNSVTLTSKKGATFQGFAGFNFGDDLVSGLWLAEGSTLQVFWQSSGILPSN 180  
QY 275 CSLPHVYVNI--KATKLSRHSVFSSYQDHAKWCISOKGTKNRWTICIGDLNRSPhQAFRS 332  
Db 191 CSGAQHVFDVQTAFPGAGAPAFNATEDHRSKWCVPKGP--WACVGDWNRNQRERHGG 237  
QY 333 GFCITQNNQIYQAFQGLVLYVESCK 357  
Db 238 GTLCAQ--MLMKAFPLVKAMEPCE 260

RESULT 9  
ABB63496  
ID ABB63496 standard; protein; 366 AA.  
XX  
AC ABB63496;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 17280.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
DR WPI; 2001-656860/75.  
DR N-PSDB; ABL07599.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions.

XX Disclosure; SEQ ID NO 17280; 21pp + Sequence Listing; English.  
PS  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABU16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 366 AA;

Query Match 23.2%; Score 449; DB 4; Length 366;  
Best Local Similarity 32.6%; Pred. NO. 1.6e-33;  
Matches 120; Conservative 68; Mismatches 148; Indels 32; Gaps 15;  
QY 10 FALLFLGLFGLVGA-ATISCTNREGKAVDVPFTFYKLPKR-QNKESGE--TGLEYLYLDST 65  
Db 6 FVLLFVFFVFFQNEAKSKVSCCKDEAGNDVDMWHLKLPKHYQHNDLGDTSGLKYLIVTSQ 65  
QY 66 T-RSWRKSEQLMNDTKSVGLGTLOOLYEAYASKNNNTAYLIYNDGVPKPNVYSRKYGHTK 124  
Db 66 NYDTWQMSGKFI SDPLSLPAQTLNPLND---DPSHTLLAAAYNDQPNNGTVFS-SGGHAK 120  
QY 125 GLLLNVRVQGLIHSIPOPPPPIPEEGYDYDPTGRRNGSGGICITFKYNOYEADISQLLV 184  
Db 121 GVASDGTATWIVHSVFKFTIPD--YSYPTSGEQVQAQSMCLCVLKGEDEKVK-QGLIV 177  
QY 185 CN-PNVYSCSIP-ATFHQELIHPQLCTRAS----SSHIPGRLITLQSAQGGKFLHFAK 238  
Db 178 YNEPHFYQRPNPLATRSDEL--FPSL-ERALHGWRTESPPQKLEVSRLDGKPKFLGK 234  
QY 239 SDSFLDDIFAAWMAQRKTHLLTETWORKQELPSNCSLPYHVYNIKAIKLSRHSY-FSS 297  
Db 235 SGRANVELYADWVAPTLDVSLFVEAWRDGAGNLPNSCDKSKVLNVESISNPELSVDFKT 294  
QY 298 YODHAKWCISQ-----KGTKNRWTICIGDLNRSPhQAFRSGGFICTONWQIYQAFQ 348  
Db 295 TQDSKWAIVSRPTGILIYHVRVGGDWICVGDINRQEGQLHRRGGTVCHKSAVSNLIRQ 354  
QY 349 LVLYYESC 356  
Db 355 LVTVYDKC 362

RESULT 10  
ABG72082  
ID ABG72082 standard; protein; 192 AA.  
XX  
AC ABG72082;  
XX  
DT 10-FEB-2003 (first entry)  
XX  
DE Mouse endonuclease implicated in switch recombination (Endo-SR).  
XX  
KW Mouse; enzyme; endonuclease implicated in switch recombination; Endo-SR;  
KW G-rich region; DNA rearrangement; Genetic recombination; B lymphocyte;  
KW Ig constant region; immune system; B cell; ligand specificity;  
KW gene therapy; tumour; cancer; leukaemia; lymphoma; apoptosis;  
KW programmed cell death; peptidomimetic.  
XX  
OS Mus sp.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 142..143  
FT /note= "Encoded by CCTACACCTATCCCTTGTCTATGACCAAGCTGGA  
FT AGGCTTCTTGCTCAGAAATACCTGAC"  
XX  
FN US6455250-B1.

XX PD 24-SEP-2002.  
 XX PF 11-DEC-1998; 98US-00210422.  
 XX PR 11-DEC-1997; 97US-0069205P.  
 XX PA (REGC ) UNIV CALIFORNIA.  
 XX PI Aguilera RJ, Lyon CJ;  
 XX DR WPI; 2003-066229/06.  
 XX DR N-PSDB; ABS58030.  
 XX PT New nucleic acid and its encoded murine endonuclease-SR protein, useful  
 XX PT for treating tumors or cancers, e.g. leukemias or lymphomas, or for  
 XX PT designing and isolating peptidomimetics or inhibitors of Endo-SR for  
 XX PT treating these diseases.  
 XX PS Example 5; Col 82; 66pp; English.  
 XX CC The invention discloses an isolated nucleic acid molecule, which encodes  
 XX CC a murine endonuclease implicated in switch recombination (Endo-SR)  
 XX CC protein. The enzyme is useful at cleaving DNA at specific G-rich regions  
 XX CC which are implicated in modulating DNA rearrangements. Also disclosed are  
 XX CC methods for repairing DNA and modulating genetic recombination in a cell.  
 XX CC One example of genetic rearrangement is in mature B lymphocytes which can  
 XX CC alter their Ig constant region for another, significantly enhancing the  
 XX CC versatility of the immune system by allowing B cells to alter their  
 XX CC function without altering their ligand specificity. The endo-SR nucleic  
 XX CC acid and protein are useful (using gene therapy) for treating tumours or  
 XX CC cancers, e.g. leukaemias or lymphomas, as well as the modulation of  
 XX CC apoptosis and programmed cell death events. The endo-SR nucleic acid and  
 XX CC protein are also useful for designing and isolating peptidomimetics and  
 XX CC inhibitors of Endo-SR, which may be employed for treating the diseases.  
 XX CC The sequence presented is the mouse endo-SR protein  
 XX SQ Sequence 192 AA;  
 Query Match 11.8%; Score 229; DB 6; Length 192;  
 Best Local Similarity 26.8%; Pred. No. 4e-13;  
 Matches 67; Conservative 25; Mismatches 95; Indels 66; Gaps 6;  
 QY 6 LRTSFALLFLGVLGALATISCRNEGKAVDWFYFKLPKQKESGEGTGLEVLVDST 65  
 Db 1 MATLSILLAALLWV-PAEALSCVGSQGVDFVYVYKLPKQKESGEGTGLEVLVDST 65  
 QY 66 TRSWRKSEQLMNDTKSVLGRITLQOLYEAVASKSNNTAYLLYNDGVPK--PWNYSRKYGHT 123  
 Db 60 SDGWQDGVGYNSPEGAVGRSLQPLR---KNSSQLAFLLYNDQPPKSSSTRDSTGHGHT 116  
 QY 124 KGLLWNRVQGFLLHISIPQFPPIPEBEGDYPTGRRNGSGCICIFKYNQYRAIDSQL 193  
 Db 117 KG-----KQLT 122  
 QY 184 VCNPNVYSCSIPATFQELIHPQLCTRASSBI---PGRLLTLOSAGOKFLHAKSD 240  
 Db 123 YTVPLVYDKLGEFFAQL---PDLETIKQHVLPWNSVILLISQAGATFQSPAKFG 179  
 QY 241 SFLLDDIFAAWMAQ 253  
 Db 180 KFGDDLYSGWLAE 192  
 RESULT 11  
 ABU41838  
 ID ABU41838 standard; protein; 391 AA.  
 XX AC ABU41838;  
 XX 19-JUN-2003 (first entry)  
 XX DE Protein encoded by Prokaryotic essential gene #27365.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 XX OS Pseudomonas syringae.  
 XX PN WO200277183-A2.  
 XX PD 03-OCT-2002.  
 XX PF 21-MAR-2002; 2002WO-US009107.  
 XX PR 21-MAR-2001; 2001US-00815242.  
 XX PR 06-SEP-2001; 2001US-00948993.  
 XX PR 25-OCT-2001; 2001US-0342923P.  
 XX PR 08-FEB-2002; 2002US-00072851.  
 XX PR 06-MAR-2002; 2002US-0362699P.  
 XX PA (ELIT-) ELITRA PHARM INC.  
 XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 XX PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;  
 XX DR WPI; 2003-029926/02.  
 XX DR N-PSDB; ACA45708.  
 XX CC New antisense nucleic acids, useful for identifying proteins or screening  
 XX CC for homologous nucleic acids required for cellular proliferation to  
 XX CC isolate candidate molecules for rational drug discovery programs.  
 XX PS Claim 25; SEQ ID NO 69762; 1766pp; English.  
 XX CC The invention relates to an isolated nucleic acid comprising any one of  
 XX CC the 6213 antisense sequences given in the specification where expression  
 XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 XX CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 XX CC encoding a polypeptide whose expression is inhibited by the antisense  
 XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 XX CC polypeptide or its fragment whose expression is inhibited by the  
 XX CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 XX CC proliferation; (7) identifying a gene in an operon required for  
 XX CC the gene product or that has an activity against a biological pathway  
 XX CC required for proliferation, or that inhibits cellular proliferation; (8)  
 XX CC identifying a gene required for cellular proliferation or the biological  
 XX CC pathway in which a proliferation-required gene or its gene product lies  
 XX CC or a gene on which the test compound that inhibits proliferation of an  
 XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 XX CC product is overexpressed or underexpressed; (12) determining the extent  
 XX CC to which each of the strains is present in a culture or collection of  
 XX CC strains; or (13) identifying the target of a compound that inhibits the  
 XX CC proliferation of an organism. The antisense nucleic acids are useful for  
 XX CC identifying proteins or screening for homologous nucleic acids required  
 XX CC for cellular proliferation to isolate candidate molecules for rational  
 XX CC drug discovery programs, or for screening homologous nucleic acids  
 XX CC required for proliferation in cells other than S. aureus, S. typhimurium,  
 XX CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of  
 XX CC the target prokaryotic essential genes. Note: The sequence data for this  
 XX CC patent did not form part of the printed specification, but was obtained  
 XX CC in electronic format directly from WIPO at  
 XX CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 391 AA;  
 Query Match 5.1%; Score 98.5; DB 6; Length 391;  
 Best Local Similarity 24.0%; Pred. No. 2.3;  
 Matches 63; Conservative 33; Mismatches 110; Indels 57; Gaps 11;  
 QY 15 LGLFVLGAATISCRNEGKAVDWFYFKLPKQKESGEGTGLEVLVDST 64  
 Db 10 LGL-GAMGAATVYQLAKAGVDVIGVDYRAPPHTQSSGSDTITRLSVGEGQYLPVRS 68





CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
xx Sequence 1272 AA;

Query Match	5.1%; Score 98.5; DB 4; Length 1272;
Best Local Similarity	17.5%; Pred. No. 12;
Matches	61; Conservative 52; Mismatches 136; Indels 99; Gaps 12;

QY	23	AATISCRNEGKAVDF-TFKLPKPKQNKSGETGLYLEYLDSTTRSWRKSEQLMNDTKS	81
DB	283	ASGITSAGDSVASWTCDYDGEDSD-----LSSTKPGSHSPSVALNPGSR	330
QY	82	VLGRTLOQLYEAASKENNTAYLIYDGVKPKVNVSRKYGHTKGLLL-----WNRVQGFMLI	138
DB	331	DL-----SSAKPSGHSRSVVLNPG-PKDISSAKPSGHSRSVVLNPGAGSAQGYSMV	380
QY	139	HSIP-----QPPIPEGYD-----YPPTGRRNGQSGIC	167
DB	381	QPSFWKQESWAQCNGOTLAVTNSQSLGWVSFPASAPGLENTKMSPT-----	430
QY	168	ITEKYNQYEALDSQLACNFENVYSCSIPATFHOELIHPOLCTRASSEETPGLLITLQS	227
DB	431	-----QLDLVTWNSVRCSAVGLHCVAPDF--RLVSFVFLTKMHGFPAGRIKNVLN	482
QY	228	AQSQKTLFHAQKSDFLDDIPFAAWMAQELKTHLLTETWQRKQELPNCSPLYHYVMIKAI	287
DB	483	PVRMLLLISFNVSFAFSQSQCWLAVRTI-----LQHLQSDYVAFALISLIVCI	532
QY	288	KLRSRHSYFSYQDHAKWCISQK-----TKNRWTCIGDLNRSPhQAFRS	331
DB	533	SF-----NHVREAVGRKNHRAHKDFWACSGGLQTCFPGKLOS	570

RESULT 14	
ABG91036	
ID	ABG91036 standard; protein; 697 AA.
XX	
AC	ABG91036;
XX	
DT	29-NOV-2002 (first entry)
XX	
DE	Chlamydia trachomatis hypothetical protein #4.
XX	
KW	Gram-negative bacterial bleb; PorB; outer membrane protein;
KW	Chlamydia trachomatis infection; Chlamydia pneumoniae infection;
KW	protective antigen; antibacterial; vaccine.
XX	
OS	Chlamydia trachomatis.
XX	
FPN	WO2000262380-A2.
XX	
XX	15-AUG-2002.
XX	
PFT	08-FEB-2002; 2002WO-EP001356.
XX	
PR	08-FEB-2001; 2001GB-00003169.
XX	
PPA	(SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX	
PI	Berthet FJ, Lobet Y, Poolman J, Verlant VGCL;
XX	
DR	WPI; 2002-657510/70.
DR	N-PSDB; ABS67357.
XX	
PFT	Novel gram-negative bacterial bleb presenting on its surface PorB outer
PFT	membrane protein from Chlamydia trachomatis or protective antigen from
PFT	Chlamydia pneumoniae, useful for preventing Chlamydia infection.
XX	

PS Disclosure; Page 20; 75pp; English.

CC The present invention relates to a new gram-negative bacterial bleb

CC presenting on its surface the PorB outer membrane protein from *Chlamydia*

CC trachomatis, or a protective antigen from *C. pneumoniae*. The invention is

CC useful for preventing *C. trachomatis* or *C. pneumoniae* infection in a

CC host. The present amino acid sequence represents a *Chlamydia trachomatis*

CC protein of the invention

CC xx

CC Sequence 697 AA:

SQ

Query Match	5.0%; Score 97.5; DB 5; Length 697;
Best Local Similarity	24.9%; Pred. No. 6.6;
Matches	62; Conservative 25; Mismatches 109; Indels 53; Gaps 12
QY	114 VNTSRK-----YCHTKGLLWN-RVQGF-WLIH----SIPQFPPIPEGVDYPPTG 158 
Db	85 VNTRGKILIDYLEYEDTDSCLLTNGRCSLYPWFIGGSTITISFSSIIHKGIYSTSEG 144 
QY	159 RRGQSGICITCFKNOYEADISOLLVNCNPVYSCISATPQHEL-----IHMPOLCWRA 212 
Db	145 ----PQKHCISGDYLVKYS--DSVLMSGFSLRISICNTPVLLPQISIMPMEIPKPPITFRG 200 
QY	213 SSSEIPGRLMTTLQSAQGFLEFAKSDSFLDDTFAAWMAQRLKTHLLTETWQRKQRLP 272 
Db	201 GSGGFLGSLGVSVSPISK--HCSTTLFLDGFKKHGIG-----LGYNRRFSSQBNP 250 
QY	273 SNGSLPVHVYNIKAIKLSRHSYFSSYODHAKWCISQKTKNRTWTCIGDLNRSPIHQAFSG 332 
Db	251 SN-----ALNIK-----SYAH-RLAIDSSGAKDRYELHGDFFSKERAHLAG 292 
QY	333 GFICTQNWQ 341 
Db	293 EFLHSDSWE 301 

RESULT 15	
ADD43730	
ID	ADD43730 standard; protein; 697 AA.
XX	
XX	AC
XX	ADD43730;
XX	
DT	15-JAN-2004 (first entry)
XX	
XX	Chlamydia trachomatis immunogenic p
DE	
XX	immunogenic; infection; Chlamydia t
KW	gene therapy; antibacterial.
XX	
XX	Chlamydia trachomatis.
OS	
XX	WO2003049762-A2.
PN	
XX	
PD	19-JUN-2003.
XX	
XX	12-DEC-2002; 2002WO-IB005761.
PF	
XX	
XX	12-DEC-2001; 2001GB-00029732.
PR	
XX	06-AUG-2002; 2002GB-00018233.
PR	
XX	14-AUG-2002; 2002GB-00018924.
XX	
XX	(CHIR-) CHIRON SPA.
XX	
XX	Grandi G, Ratti G;
PI	
XX	WPI; 2003-532882/50.
XX	N-PSDB; ADD43731.
XX	
PPT	New immunogenic composition having a
PPT	useful for diagnosing, preventing an
PPT	infection.
XX	
XX	Claim 6; SEQ ID NO 25; 164pp; Enclis

Fri Oct 15 10:01:06 2004

XX

Query Match 5.0%; Score 97.5; DB 7; Length 697;

Best Local Similarity	24.9%;	Pred. No. 6.6;	12;
Matches	62;	Conservative	25;
Mismatches	25;	Mismatches	109;
Indels	53;	Gaps	12;

nb 85 VNYRGKILICDYLEYVEDTDSCLLTNGRCSLYPWFIFGGSTITISPSIIHKGYISTSEG 144

[illegible]

QY

QY 273 SNCSLPYHVYNIKAIKLSRHSYFSSYQDHAQWICISQGTKNRWTCTIGDLNRSPHQAFRSG 332

Qv 333 GFICTONWO 341

Search completed: October 14, 2004, 00:28:08  
Job time : 73.8017 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2004, 00:24:00 ; Search time 31.6329 Seconds  
(without alignments)  
582.636 Million cell updates/sec

Title: US-10-790-589-4

Perfect score: 1933

Sequence: 1 MWARLRTSFALLFLGLFGV.....QNWIYQAFGLVLYESCK 357

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/prodata/2/iaa/5A COMB.pdp.\*
- 2: /cgn2\_6/prodata/2/iaa/5B COMB.pdp.\*
- 3: /cgn2\_6/prodata/2/iaa/6A COMB.pdp.\*
- 4: /cgn2\_6/prodata/2/iaa/6B COMB.pdp.\*
- 5: /cgn2\_6/prodata/2/iaa/PTUS COMB.pdp.\*
- 6: /cgn2\_6/prodata/2/iaa/backfiles1.pdp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1933	100.0	357	4	US-09-574-942-4
2	1926	99.6	361	4	US-09-807-784B-3
3	1293.5	66.9	354	4	US-09-574-942-2
4	1293.5	66.9	354	4	US-09-807-784B-1
5	613	31.7	360	3	US-09-147-915-3
6	613	31.7	360	4	US-08-639-294-2
7	613	31.7	360	4	US-09-861-034B-2
8	458	23.7	275	3	US-09-147-915-4
9	229	11.8	192	4	US-09-210-422-2
10	96.5	5.0	661	4	US-09-198-452A-36
11	95.5	4.9	426	2	US-08-416-870C-10
12	95.5	4.9	595	2	US-08-468-036-3
13	95	4.9	595	2	US-08-376-843-3
14	95	4.9	704	4	US-09-409-180A-1
15	90.5	4.7	384	4	US-09-071-035-276
16	90.5	4.7	430	4	US-09-071-035-274
17	90.5	4.7	2037	4	US-09-543-681A-5538
18	89.5	4.6	477	4	US-09-134-000C-6123
19	88.5	4.6	542	4	US-09-543-681A-6775
20	88.5	4.6	1456	4	US-09-976-594-168
21	88	4.6	459	1	US-08-220-151-12
22	88	4.6	459	1	US-08-220-151-14
23	88	4.6	459	1	US-08-413-118-12
24	88	4.6	459	1	US-08-413-118-14
25	88	4.6	459	3	US-08-413-446-12
26	88	4.6	459	3	US-08-473-446-14
27	88	4.6	459	3	US-09-213-053-6

```

28      88      4.6      1394      3      US-09-213-053-2      Sequence 2, Appli
29      84.5      4.4      967      4      US-09-489-039A-8766      Sequence 8766, Ap
30      84.5      4.4      1455      3      US-08-840-062-5      Sequence 5, Appli
31      84      4.3      455      3      US-09-221-235-5      Sequence 5, Appli
32      84      4.3      455      3      US-09-221-928-5      Sequence 5, Appli
33      84      4.3      455      3      US-09-221-527-5      Sequence 5, Appli
34      84      4.3      455      3      US-09-221-236-5      Sequence 5, Appli
35      84      4.3      455      3      US-09-221-416-5      Sequence 5, Appli
36      84      4.3      455      3      US-09-221-245-5      Sequence 5, Appli
37      84      4.3      455      3      US-09-163-115-5      Sequence 5, Appli
38      84      4.3      455      3      US-09-221-528-5      Sequence 5, Appli
39      84      4.3      455      3      US-09-593-553-5      Sequence 5, Appli
40      84      4.3      455      3      US-09-221-237-5      Sequence 5, Appli
41      84      4.3      455      4      US-09-399-588-2      Sequence 2, Appli
42      84      4.3      528      4      US-09-328-352-6385      Sequence 6385, Ap
43      84      4.3      855      2      US-08-482-090-12      Sequence 12, Appli
44      84      4.3      856      2      US-08-481-700B-8      Sequence 8, Appli
45      84      4.3      856      2      US-09-007-383-16      Sequence 16, Appli

```

## ALIGNMENTS

### RESULT 1

US-09-574-942-4

; Sequence 4, Application US/09574942

; Patent No. 6358723

; GENERAL INFORMATION:

; APPLICANT: Eastman, Alan R.

; APPLICANT: Krieser, Ronald J.

; TITLE OF INVENTION: DROXYRIBONUCLEASE II BETA PROTEINS AND CDNAS

; FILE REFERENCE: DC-0137

; CURRENT APPLICATION NUMBER: US/09/574,942

; CURRENT FILING DATE: 2000-05-19

; PRIOR APPLICATION NUMBER: 09/541,840

; PRIOR FILING DATE: 2000-04-03

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 357

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-574-942-4

Query Match 100.0%; Score 1933; DB 4; Length 357;

Best Local similarity 100.0%; Pred. No. 1.5e-193;

Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 MWARLRTSFALLFLGLFGVLCATISCNERKGVADWFTFYKLPKRNKESGCTGLVYL 60
Db      1 MWARLRTSFALLFLGLFGVLCATISCNERKGVADWFTFYKLPKRNKESGCTGLVYL 60

QY      61 YLDSTTRGWRKSEQLMNDTKSVLGRTLQQLYEAYASKNNNTAYLYNDGVGPKPVNSRYK 120
Db      61 YLDSTTRGWRKSEQLMNDTKSVLGRTLQQLYEAYASKNNNTAYLYNDGVGPKPVNSRYK 120

QY      121 GHTKGLLLNNRVQGVFWLHISIPQFPPIPEGDYDPTGRNGQSGICITFKYNQVEAIDS 180
Db      121 GHTKGLLLNNRVQGVFWLHISIPQFPPIPEGDYDPTGRNGQSGICITFKYNQVEAIDS 180

QY      181 QLLVCNPNVYCSIPATFHQLIHPQLCTRASSSEIPGRLLITLQSAQOGKFLHFKSD 240
Db      181 QLLVCNPNVYCSIPATFHQLIHPQLCTRASSSEIPGRLLITLQSAQOGKFLHFKSD 240

QY      241 SFLDDIFAAMWAQRLLKTHLLTETWQRKQELPNSCLPQVHVNIKAIKLSRHSYSSYQD 300
Db      241 SFLDDIFAAMWAQRLLKTHLLTETWQRKQELPNSCLPQVHVNIKAIKLSRHSYSSYQD 300

QY      301 HAKWCISOKGTNRWTCIGDLNRSHPQAFRSGGFTCTQNWQIYQAFQGLVLYESCK 357
Db      301 HAKWCISOKGTNRWTCIGDLNRSHPQAFRSGGFTCTQNWQIYQAFQGLVLYESCK 357

```

```

RESULT 2
US-09-807-784B-3
; Sequence 3, Application US/09807784B
; Patent No. 6653118
; GENERAL INFORMATION:
; APPLICANT: Tanuma, Sei-ichi
; APPLICANT: Shiokawa, Daisuke
; TITLE OF INVENTION: No. 6653118el Deoxyribonuclease, Gene Encoding Thereof and Use Th
; FILE REFERENCE: 210792
; CURRENT APPLICATION NUMBER: US/09/807,784B
; CURRENT FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: JP 11-230870
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 15
; SEQ ID NO 3
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-807-784B-3

Query Match          99.6%; Score 1926; DB 4; Length 361;
Best Local Similarity 99.7%; Pred. No. 8.4e-193;
Matches 356; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MMARLLRTSALLFLGLFGVLGAATISCRNEEGKAVDWFYKLPKRONKESGETGLEYL 60
Db 5 MMARLLRTSALLFLGLFGVLGAATISCRNEEGKAVDWFYKLPKRONKESGETGLEYL 64

Qy 61 YLDSTTRSWRKSEQLMNDTKSVLGRITLQQLYEAYASKSNNTAYLIYNDGVPKPNYSRKY 120
Db 65 YLDSTTRSWRKSEQLMNDTKSVLGRITLQQLYEAYASKSNNTAYLIYNDGVPKPNYSRKY 124

Qy 121 GHTYGLLLNNRVQGFWLHISIPQPPPIPEGYDYPPTGRRNGQSGICITFKYNOYEIDS 180
Db 125 GHTYGLLLNNRVQGFWLHISIPQPPPIPEGYDYPPTGRRNGQSGICITFKYNOYEIDS 184

Qy 181 QLLVCNPNVYSCSIPATPHOELHMPOLCTRASSEIPGRLLTTLSQAQOKFLHFAKSD 240
Db 185 QLLVCNPNVYSCSIPATPHOELHMPOLCTRASSEIPGRLLTTLSQAQOKFLHFAKSD 244

Qy 241 SFLLDDIPAAWMAOQLKTHLLTETWQRKQELPNSCLPYHYVNIKAIKLSRHSYFSSYQD 300
Db 245 SFLLGIPAAWMAOQLKTHLLTETWQRKQELPNSCLPYHYVNIKAIKLSRHSYFSSYQD 304

Qy 301 HAKWCISQGTKNRWTCIGDLNRSPhQAFRSGGFICTQNWQIYQAFQGLVLYYESCK 357
Db 305 HAKWCISQGTKNRWTCIGDLNRSPhQAFRSGGFICTQNWQIYQAFQGLVLYYESCK 361

RESULT 3
US-09-574-942-2
; Sequence 2, Application US/09574942
; Patent No. 6358723
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Krieser, Ronald J.
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCE: DC-0137
; CURRENT APPLICATION NUMBER: US/09/574,942
; CURRENT FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-574-942-2

Query Match          66.9%; Score 1293.5; DB 4; Length 354;
Best Local Similarity 66.4%; Pred. No. 1.1e-126;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

```

```

Qy 1 MMARLLRTSALLFLGLFGVLGAATISCRNEEGKAVDWFYKLPKRONKESGETGLEYL 60
Db 1 MTAKPLRTVLSLFFALSGVLGPEISCRNEYEAVDWFYKLPKRTSKASEAGLQYL 60

Qy 61 YLDSTTRSWRKSEQLMNDTKSVLGRITLQQLYEAYASKSNNTAYLIYNDGVPKPNYSRKY 120
Db 61 YLDSTTRQWNKSLYLINSTRSALGRTLQHLDYTHNS--TNDTAYLIYNDGVPGSVNSRQY 119

Qy 121 GHTYGLLLNNRVQGFWLHISIPQPPPIPEGYDYPPTGRRNGQSGICITFKYNOYEIDS 180
Db 120 GHAKGLLVNWRNRTQGFWLHISVPKPPV--HGYYPTSGRRYGTGICITFGYQFEIDF 177

Qy 181 QLLVCNPNVYSCSIPATPHOELHMPOLCTRASSEIPGRLLTTLSQAQOKFLHFAKSD 240
Db 178 QLLVLQPNYISCFIPSTFHWKLIYMPRCANSSSLKIPVRYLAELHSAQGLNFVHFAKSS 237

Qy 241 SFLLDDIPAAWMAOQLKTHLLTETWQRKQELPNSCLPYHYVNIKAIKLSRHSYFSSYQD 300
Db 238 FYTDDIFTGWIAQKLTLLAQTWQKKQELPNSCLPYHYVNIKSIGVTSKSYFSSRQD 297

Qy 301 HAKWCISQGTKNRWTCIGDLNRSPhQAFRSGGFICTQNWQIYQAFQGLVLYYESCK 357
Db 298 HSKWCVSIGKSANRWTCIGDLNRSPhQAALRGGGFICTKNHYIYQAFHKLRYLRYGFC 354

RESULT 4
US-09-807-784B-1
; Sequence 1, Application US/09807784B
; Patent No. 6653118
; GENERAL INFORMATION:
; APPLICANT: Tanuma, Sei-ichi
; APPLICANT: Shiokawa, Daisuke
; TITLE OF INVENTION: No. 6653118el Deoxyribonuclease, Gene Encoding Thereof and Use The
; FILE REFERENCE: 210792
; CURRENT APPLICATION NUMBER: US/09/807,784B
; CURRENT FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: JP 11-230870
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 15
; SEQ ID NO 1
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-807-784B-1

Query Match          66.9%; Score 1293.5; DB 4; Length 354;
Best Local Similarity 66.4%; Pred. No. 1.1e-126;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

Qy 1 MMARLLRTSALLFLGLFGVLGAATISCRNEEGKAVDWFYKLPKRONKESGETGLEYL 60
Db 1 MTAKPLRTVLSLFFALSGVLGPEISCRNEYEAVDWFYKLPKRTSKASEAGLQYL 60

Qy 61 YLDSTTRSWRKSEQLMNDTKSVLGRITLQQLYEAYASKSNNTAYLIYNDGVPKPNYSRKY 120
Db 61 YLDSTRTQWNKSLYLINSTRSALGRTLQHLDYTHNS--TNDTAYLIYNDGVPGSVNSRQY 119

Qy 121 GHTYGLLLNNRVQGFWLHISIPQPPPIPEGYDYPPTGRRNGQSGICITFKYNOYEIDS 180
Db 120 GHAKGLLVNWRNRTQGFWLHISVPKPPV--HGYYPTSGRRYGTGICITFGYQFEIDF 177

Qy 181 QLLVCNPNVYSCSIPATPHOELHMPOLCTRASSEIPGRLLTTLSQAQOKFLHFAKSD 240
Db 178 QLLVLQPNYISCFIPSTFHWKLIYMPRCANSSSLKIPVRYLAELHSAQGLNFVHFAKSS 237

Qy 241 SFLLDDIPAAWMAOQLKTHLLTETWQRKQELPNSCLPYHYVNIKAIKLSRHSYFSSYQD 300
Db 238 FYTDDIFTGWIAQKLTLLAQTWQKKQELPNSCLPYHYVNIKSIGVTSKSYFSSRQD 297

Qy 301 HAKWCISQGTKNRWTCIGDLNRSPhQAFRSGGFICTQNWQIYQAFQGLVLYYESCK 357
Db 298 HSKWCVSIGKSANRWTCIGDLNRSPhQAALRGGGFICTKNHYIYQAFHKLRYLRYGFC 354

```

## RESULT 5

US-09-147-915-3  
; Sequence 3, Application US/09147915A  
; Patent No. 6184034  
; GENERAL INFORMATION:  
; APPLICANT: Eastman, Alan  
; APPLICANT: Krieser, Ronald  
; TITLE OF INVENTION: Deoxyribonuclease II Proteins and cDNAs  
; FILE REFERENCE: DC-0097  
; CURRENT APPLICATION NUMBER: US/09/147,915A  
; CURRENT FILING DATE: 1999-03-23  
; EARLIER APPLICATION NUMBER: PCT/US97/18262  
; EARLIER FILING DATE: 1997-10-09  
; EARLIER APPLICATION NUMBER: 60/028,539  
; EARLIER FILING DATE: 1996-10-15  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 360  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-147-915-3

Query Match 31.7%; Score 613; DB 3; Length 360;

Best Local Similarity 36.9%; Pred. No. 1.7e-55;  
Matches 129; Conservative 61; Mismatches 148; Indels 12; Gaps 6;

QY 13 LFLGLFVLGAATISCRNEEGKAVDWFTFYKLPK-RONKESGETGLLEYLYLDSTTSRWK 71  
Db 4 LLLAALLCPAGALTCYGDGQPDWVFVYKLPALRGSGEAAQGRQYKYLDSSGGWRD 63  
QY 72 SSQLMNDTKSVLGRITLQOLYEAVASKNNTAVLYINDGVKPK--VNYSRKYGHTKGLLLW 129  
Db 64 GRALINSPGAVGRSLQPLYR---SNTSOLAFLYNDQPPQSKAQDSSMRGHTKGVLL 120  
QY 130 NRVOGFWLHISIPQF--PTPIPERGYDYPPTGRNGSGGICITFKYNOYEAIDSQLLVCNPN 188  
Db 121 DHGQGFVLVHSPNPPASSAAAYSWPHSACTYGTQLLCVSPFPAQFSKMGKQLTYTYPW 180  
QY 189 VYSCIPATFHQELHMPOLCTRASSEIPGRLLTTLQSAQOQKFLHFAKSDSFLDDIFA 248  
Db 181 VVNYOLEGIFAQEFPPDLENVVKGHVSPQPMNSSITLTSQAGAVFQSFQKFKFGDDLYS 240  
QY 249 ANWAQRLKTHLTETWQRKQELPNSCLPYHYVNAIKALKSRHS--YFSSYQDHAKWCI 306  
Db 241 GWLAALGNLQVQVWHKTVGILPNSCDIWQVNLVNIQIAFFPGAPGSPFNSTEDHSKNCV 300  
QY 307 SOKGTNRWTCIGDLNRSHPQAFRSGGFTCTQNWQIYQAFQGLVLYYESC 356  
Db 301 SPKGP---WTCVGMNRNQGEGRGGTTLCAQLPALWKAQFLVKNYQPC 347

## RESULT 6

US-08-639-294-2  
; Sequence 2, Application US/08639294  
; Patent No. 6265195  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Baron, Will F.  
; TITLE OF INVENTION: Human DNase II  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/639,294  
; FILING DATE: 25-Apr-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Johnston, Sean A.  
; REGISTRATION NUMBER: 35,910  
; REFERENCE/DOCKET NUMBER: P1024  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-3562  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 360 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
US-08-639-294-2

Query Match 31.7%; Score 613; DB 3; Length 360;

Best Local Similarity 36.9%; Pred. No. 1.7e-55;  
Matches 129; Conservative 61; Mismatches 148; Indels 12; Gaps 6;

QY 13 LFLGLFVLGAATISCRNEEGKAVDWFTFYKLPK-RONKESGETGLLEYLYLDSTTSRWK 71  
Db 4 LLLAALLCPAGALTCYGDGQPDWVFVYKLPALRGSGEAAQGRQYKYLDSSGGWRD 63  
QY 72 SSQLMNDTKSVLGRITLQOLYEAVASKNNTAVLYINDGVKPK--VNYSRKYGHTKGLLLW 129  
Db 64 GRALINSPGAVGRSLQPLYR---SNTSOLAFLYNDQPPQSKAQDSSMRGHTKGVLL 120  
QY 130 NRVOGFWLHISIPQF--PTPIPERGYDYPPTGRNGSGGICITFKYNOYEAIDSQLLVCNPN 188  
Db 121 DHGQGFVLVHSPNPPASSAAAYSWPHSACTYGTQLLCVSPFPAQFSKMGKQLTYTYPW 180  
QY 189 VYSCIPATFHQELHMPOLCTRASSEIPGRLLTTLQSAQOQKFLHFAKSDSFLDDIFA 248  
Db 181 VVNYOLEGIFAQEFPPDLENVVKGHVSPQPMNSSITLTSQAGAVFQSFQKFKFGDDLYS 240  
QY 249 ANWAQRLKTHLTETWQRKQELPNSCLPYHYVNAIKALKSRHS--YFSSYQDHAKWCI 306  
Db 241 GWLAALGNLQVQVWHKTVGILPNSCDIWQVNLVNIQIAFFPGAPGSPFNSTEDHSKNCV 300  
QY 307 SOKGTNRWTCIGDLNRSHPQAFRSGGFTCTQNWQIYQAFQGLVLYYESC 356  
Db 301 SPKGP---WTCVGMNRNQGEGRGGTTLCAQLPALWKAQFLVKNYQPC 347

## RESULT 7

US-08-661-034B-2  
; Sequence 2, Application US/09861034B  
; Patent No. 6569429  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Baron, Will F.  
; TITLE OF INVENTION: Human DNase II  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/861,034B

```

; FILING DATE: 18-May-2001
; CLASSIFICATION: <UNKNOWN>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/639294
; FILING DATE: 25-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Johnston, Sean A.
; REGISTRATION NUMBER: 35,910
; REFERENCE/DOCKET NUMBER: P1024D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3562
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-861-034B-2

```

```

Query Match 31.7%; Score 613; DB 4; Length 360;
Best Local Similarity 36.9%; Pred. No. 1.7e-55;
Matches 129; Conservative 61; Mismatches 148; Indels 12; Gaps 6;

QY 13 LFLGLFGLVGAATISCRNEEGKAVDFTFYKLPK-RONKESGETGLLEYLYLDSTTRGWRK 71
DB 4 LLLAALCVFAGALTCTGDSGQPDVDFVVKLPALRGSGEAAQGLQYKYLDESSEGGWRD 63

QY 72 SEQLMNDTKSVGLRTLOQLYAYASKNNATLYINDGVKPK--VNYSRKYGHTKGLLLW 129
DB 64 GRALINSEGAAGVGRSLQPLYR--SNTSQLAFLYNDQPPQPSKAQDSMRGHTKGVILL 120

QY 130 NRVOGFWLIHSIOP-PRPEEGDYPTGRNGOSGICITFKYNOYEADISOLLVCNPN 188
DB 121 DHGDFGLVHSPNFPFPASSAAYSWPHSACTYQTLTLCVSPFAQSKMGKQLYTYVPW 180

QY 189 VYSCISIPATHOELIHPOLCTRASSSEIPRLMTLQSAQGLFLHFAKSDSFLDDIFA 248
DB 181 VYVQLSGIAFQEPDLENVVKGHVSGEPWNSSITLTSQAGAVFQSPFAKESKFGDLYS 240

QY 249 AWAORLKLHLLTWTQKROKQLPNSCLPYPHYVNIKAIKLSRHS--YFSSYQDHAKWCI 306
DB 241 GWLAAALGTNLQVQFHKTVGILPNSGSDIWOVLNVNQLAPPGAPGPFSEFSTEDHSKWC 300

QY 307 SOKGTNRWTCIGDNRSPHQAQFSGGFICTQNWQIYCAQFGLVLYYESC 356
DB 301 SPKGP---WTCVGMNRNQGEEQGGGTLCALPALWKAFOPLVKNYQPC 347

```

```

RESULT 8
US-09-147-915-4
; Sequence 4, Application US/09147915A
; Patent No. 6184034
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan
; APPLICANT: Krieser, Ronald
; TITLE OF INVENTION: Deoxyribonuclease II Proteins and cDNAs
; FILE REFERENCE: DC-0097
; CURRENT APPLICATION NUMBER: US/09/147,915A
; CURRENT FILING DATE: 1999-03-23
; EARLIER APPLICATION NUMBER: PCT/US97/18262
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/028,539
; EARLIER FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 4
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Bos sp.
US-09-147-915-4

```

Query Match 23.7%; Score 458; DB 3; Length 275;

```

Best Local Similarity 36.4%; Pred. No. 1.8e-39;
Matches 95; Conservative 50; Mismatches 106; Indels 10; Gaps 5;

QY 102 AYLIYNDGPKFVNY--SRKYGHTKGLLWNRVQGWFLHISIPQFP-IFEEGVYDVPPTG 158
DB 4 AFVLYNDQPPKSESXSSRSGHTKGVLLDQSGFWLHISVFNFPASSAAYSWPPGA 63

QY 159 RRGOSGICITFKYNOYEADISOLLVCNPNVYSCISIPATHOELIHPOLCTRASSSEIP 218
DB 64 OKTGTLLICVSPFLTQFLDISKQLYTYPLVYDHRLEGDFGQFPFVLEEVKGHVQGP 123

QY 219 GRLLTTLQSAQGLFLHFAKSDSFLDDIFAAMQAORLKLHLLTETWORKQELPNSCLP 278
DB 124 WNSSVTLTSSKGTATFQSFQAKFGNFGDLYSGWLAALGSLTQVQFWORSSGILPNSCSGA 183

QY 279 YHYVNI--KAIKLSRHSYFSSYQDHAKWCI-SOKGTNRWTCIGDNRSPHQAQFSGGFI 336
DB 184 QHVFDTVTQTAPPGAPGAFNATEDHSKWCVTPKGP---WACVGMNRNQRREHRRGGTLC 240

QY 337 TONWQIYCAQFGLVLYYESC 357
DB 241 AQ--MLWKAFKPLVKAWEPC 259

```

```

RESULT 9
US-09-210-422-2
; Sequence 2, Application US/09210422A
; Patent No. 6455250
; GENERAL INFORMATION:
; APPLICANT: AGUILERA, RENATO J.
; APPLICANT: LYON, CHRISTOPHER J.
; TITLE OF INVENTION: ENDONUCLEASE COMPOSITIONS AND METHODS OF USE
; FILE REFERENCE: UCLA019
; CURRENT APPLICATION NUMBER: US/09/210,422A
; CURRENT FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 60/069,205
; PRIOR FILING DATE: 1997-12-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Musca sp.
US-09-210-422-2

```

```

Query Match 11.8%; Score 229; DB 4; Length 192;
Best Local Similarity 26.5%; Pred. No. 9.1e-16;
Matches 67; Conservative 25; Mismatches 95; Indels 66; Gaps 6;

QY 6 LRTSFALLFLGLFGLVGAATISCRNEEGKAVDFTFYKLPKRONKESGETGLLEYLYLDST 65
DB 1 MATLRSLLLAALLWV-PAEALSCYDGSQPDVDFVVKLPALHSGSRDTPKGLTCKYMDQ 59

QY 66 TRSWRKSEQLMNDTKSVGLRTLOQLYAYASKNNATLYINDGVKPK--PVNYSRKYGH 123
DB 60 SDGWQDGVGVINSPEGAAGVGRSLQPLYR---KNSSQLAFLYNDQPPKSSSTRDSTGH 116

QY 124 KGLLWNRVQGLHISIPQFPPIPEEGDYPTGRNGOSGICITFKYNOYEADISOLL 183
DB 117 KG-----KQLT 122

QY 184 VCNPNVYSCISIPATHOELIHPOLCTRASSSEI---PGRLLTTLQSAQGLFLHFAKSD 240
DB 123 YTYPLVYDHRLEGDFGQAKL---PDLETWIKQHVLPWNSSVLTLSQAGATFQSFQAKFG 179

QY 241 SFLDDIFAAMQAQ 253
DB 180 KFGDDLYSGWLA 192

```

```

RESULT 10
US-09-198-452A-36
; Sequence 36, Application US/09198452A

```

```

; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griflais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 36
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-36

```

Query Match 5.0%; Score 96.5; DB 4; Length 661;  
Best Local Similarity 19.7%; Pred. No. 0.45;  
Matches 76; Conservative 59; Mismatches 157; Indels 93; Gaps 19;

QY	4	RLRLTGFALLFLGLFGVLGAATISCRNEEG--KAVDMFTFKLPRKQNKESGETGLB--Y	59
Dd	9	RLMKRCF--LFLASFLVMGSSADALTHQEAVKKNSYLSHFK-----SVSGVITIEDGV	60
QY	60	LYLDSITRSWRKBEQLMNDTKSVLGRTLQOLYEA-----YASKNNPAYLLIYNDGVKPV	114
Dd	61	LNTHNNLRIQANKVYYEN-----TVGOSLKLVAHGVNMVNYRAKTLVCDILEY-----	108
QY	115	NYSRKYGHYKGLLLWN--RVQGFLLHSPQPPP---TPEEGYDPPTRRRNGSGIGI	168
Dd	109	-----YEDTDSCLLTNGRFAMYFWFJGGSMITUTPETIVIRKGYISTSEBPK---KDLCL	160
QY	169	TFKYQYEAIDSQLLYCNPVNSCSPATPFQBEL-----IHPQLCTRASSEIPERLL	222
Dd	161	SGDYLEYSS--DSLISGKTLRLVCRIPIILFLPPFSIMPMELPKPPINFRGCTGCGFLGSYL	219
QY	223	TTLQSAQGQKFLHFAKDSFLDDIPAAWMAQRUKTHLLTETWORKRQELPSNCSLPYHVY	282
Dd	220	GMSYGPISRK--HFS--STFLDLSFFKHVGWGMGNLHC-----SQXQVPENVF	263
QY	283	NIKA-----IKLSRHSYFSSODHAKWCISQKGT-----KNRWTCIGDL-----	321
Dd	264	NMKSYAHELADMAEAHDORYLHGDFCTTHKHVNFSGEYHLSDSWETVADIPFNNFMILK	323
QY	322	NRSPHOAFPSGGFICTQNWQIYQAF	346
Dd	324	NTGPTRV-----DCTWNDNYFEGY	342

```

RESULT 11
US-08-416-870C-10
; Sequence 10, Application US/08416870C-10
; Patent No. 5824862
; GENERAL INFORMATION:
; APPLICANT: HIYOSHI, TORU
; APPLICANT: MINE, TOSHIKI
; APPLICANT: KASAKURA, KEITSUKE
; APPLICANT: TYSON, ROBERT HOW
; APPLICANT: PAGE, ANTHONY MILES JOE
; TITLE OF INVENTION: DNA ENCODING A
; TITLE OF INVENTION: 1-PHOSPHOTRANS
; TITLE OF INVENTION: SAME AND METHOD
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESSES:
; ADDRESS: BIRCH, STEWART, KOLAC
; STREET: PO BOX 747
; CITY: FALL CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,870C
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY JR, GERALD M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 760-195P(PCT)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 426 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-416-870C-10

Query Match 4.9%; Score 95.5; DB 2; Length 426;
Best Local Similarity 18.2%; Pred. No. 0.29;
Matches 67; Conservative 52; Mismatches 100; Indels 149; Gaps 20;

Qy      . 65 TTSWRKSEQLMNDTKSVLGRTLQOLYEAVYAKSNNTAYLIYND-----GYPKP 113
Db      . 116 TSRGHDHTTKIVD---SIQDRGINQVYIIIGDGSQKGAAVIFEIRRRGLKVAVAGIPKT 172

Qy      . 114 VN-----YSRKYG-----HTK-----GLL-LWNRVQGFWLHS----- 140
Db      . 173 IDNDIPIIDRSFGFDTAVEEAQRAINAAHVEATSFENGIGLVLMGRYSGFIAMATLAS 232

Qy      . 141 -----IPOFPPIPERGYDPTGRRNGOSGICITPKYQYEAIDSQLLVCNPNVYSCS 193
Db      . 233 RDVDCCLIPESPFFLE-----GKGGL-----PEFIGREL----- 261

Qy      . 194 IPATFHQELHMQLCTRASSEI-----PGRLLTTLSAQQKFLHFAKSDSFLLDIPAAW 250
Db      . 262 -----KEIGHMIVIAEGAGQDLLAESNEQSTTLKDSGNK-----LLQDV-GLW 305

Qy      . 251 MAORLTKHLITETWQRK-----QELPSN-----CSL-----PYHVTNIK 285
Db      . 306 ISQRIKDHFAKWTMLNKYIDPTMYIRAVESNADNVCCITLAAQSAVHYGMAGNGFTVG 365

Qy      . 286 AIKLSRHSYFSSYQDHAKWCISQKGTK-----NRWTCTIGDLNRSF-----HOAFRS 332
Db      . 366 LVN-GRHTYIPFYR-----ITEKQKVITDRMARLLSSTNQPSFMKDDHHPNHS 418

Qy      . 333 GFCTQNW 340
Db      . 419 GEAGAMNW 426

```

```

RESULT 12
US-08-468-036-3
; Sequence 3, Application US/08468036
; Patent No. 5728806
; GENERAL INFORMATION:
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Hoekstra, Merl F.
; TITLE OF INVENTION: Materials and Methods Relating to Proteins that
; TITLE OF INVENTION: Interact with Casein Kinase I
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,036
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/184,605
FILING DATE: 21-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: No. 5728806and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/31784
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 595 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-468-036-3

Query Match 4.9%; Score 95; DB 1; Length 595;
Best Local Similarity 22.5%; Pred. No. 0.55;
Matches 64; Conservative 35; Mismatches 99; Indels 86; Gaps 15;

QY 17 LFGVLGATII-----SCRNEGKAVDWFTFYKLPKQNKESGETGLELYLDSTTR---S 68
DB 179 IFLISGVSEVFNFKPQVRNEIDK-IEWDFKISKITWK-----SNIKYILNSMWRPLSM 233
QY 69 W-RKSEQIAMD-----TKSVLGR-----LQOQYEAAYASKNNAYLIY 106
DB 234 WLRHQRIKNEIDQLKSYAEQKLLGTTKEEQIDPGRELINMLHTAVQANSNNA--VS 291
QY 107 NDGVPKP-----VNSRYKGYHTKGLL----- 127
DB 292 NGQVPSSEQLQHLKEQSGEHNQKQDQSSFSQQQPSIFPSSLRPFANKNVIPPTMPMA 351
QY 128 --LNNRVQGFWLIHSIPQFPPIPEGDYDPTGRRNGOSGICITPKYNOYEAIDSLVLC 185
DB 352 NVFMSNPOLFATMGQP-FAPFP---FMLPLTNNNSANPIPTVPPN-FNAPNPMFAG 406
QY 186 NPNVYSCSIPATFHQELIHPQLCTRAS--SSEIPGRLLTTLOS 227
DB 407 VPMNHLGSPAV-SQPFSLPAPLPRDSGYSSSPGQLLDILNS 449

RESULT 13
US-08-376-843-3
; Sequence 3, Application US/08376843
; Patent No. 5846764
; GENERAL INFORMATION:
; APPLICANT: Demaggio, Anthony J.
; APPLICANT: Hoeckstra, Merl F.
; TITLE OF INVENTION: Materials and Methods Relating to Proteins
; TITLE OF INVENTION: that Interact with Casein Kinase I
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/376,843
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/184,605
FILING DATE: 21-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: No. 5846764and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/31784
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 595 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-376-843-3

Query Match 4.9%; Score 95; DB 2; Length 595;
Best Local Similarity 22.5%; Pred. No. 0.55;
Matches 64; Conservative 35; Mismatches 99; Indels 86; Gaps 15;

QY 17 LFGVLGATII-----SCRNEGKAVDWFTFYKLPKQNKESGETGLELYLDSTTR---S 68
DB 179 IFLISGVSEVFNFKPQVRNEIDK-IEWDFKISKITWK-----SNIKYILNSMWRPLSM 233
QY 69 W-RKSEQIAMD-----TKSVLGR-----LQOQYEAAYASKNNAYLIY 106
DB 234 WLRHQRIKNEIDQLKSYAEQKLLGTTKEEQIDPGRELINMLHTAVQANSNNA--VS 291
QY 107 NDGVPKP-----VNSRYKGYHTKGLL----- 127
DB 292 NGQVPSSEQLQHLKEQSGEHNQKQDQSSFSQQQPSIFPSSLRPFANKNVIPPTMPMA 351
QY 128 --LNNRVQGFWLIHSIPQFPPIPEGDYDPTGRRNGOSGICITPKYNOYEAIDSLVLC 185
DB 352 NVFMSNPOLFATMGQP-FAPFP---FMLPLTNNNSANPIPTVPPN-FNAPNPMFAG 406
QY 186 NPNVYSCSIPATFHQELIHPQLCTRAS--SSEIPGRLLTTLOS 227
DB 407 VPMNHLGSPAV-SQPFSLPAPLPRDSGYSSSPGQLLDILNS 449

RESULT 14
US-09-409-180A-1
; Sequence 1, Application US/09409180A
; Patent No. 6444802
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: White, David
; APPLICANT: Silos-Santiago, Inmaculada
; TITLE OF INVENTION: 22196, A No. 6444802el Human Amino peptidase
; FILE REFERENCE: 5800-59
; CURRENT APPLICATION NUMBER: US/09/409,180A
; CURRENT FILING DATE: 1999-09-30
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 704
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-409-180A-1

Query Match 4.9%; Score 95; DB 4; Length 704;
Best Local Similarity 18.7%; Pred. No. 0.71;
Matches 61; Conservative 49; Mismatches 117; Indels 100; Gaps 14;

QY 20 VLGAATISCRNE-----EGKAVDWFTFYKLPKQNKES---SGETGLELYLDSTTRSWRKS 72
```





***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2004, 00:31:21 ; Search time 67.7848 Seconds  
(without alignments)  
1694.809 Million cell updates/sec

Title: US-10-790-589-4  
Perfect score: 1933  
Sequence: 1 MMARLRTSFALLFLGLFGV.....QNWQIYQAFQGLVLYESCK 357

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pcp:  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pcp:  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pcp:  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pcp:  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pcp:  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pcp:  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pcp:  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pcp:  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pcp:  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pcp:  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pcp:  
12: /cgn2\_6/ptodata/1/pubpaa/US09D\_PUBCOMB.pcp:  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pcp:  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pcp:  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pcp:  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pcp:  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pcp:  
18: /cgn2\_6/ptodata/1/pubpaa/US10F\_PUBCOMB.pcp:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	1933	100.0	357	US-09-949-434-4	Sequence 4, Appli
2	1933	100.0	357	US-10-240-709-4	Sequence 4, Appli
3	1933	100.0	357	US-10-790-589-4	Sequence 4, Appli
4	1293.5	99.6	361	US-10-670-863-3	Sequence 2, Appli
5	1293.5	66.9	354	US-09-949-434-2	Sequence 3, Appli
6	1293.5	66.9	354	US-10-240-709-2	Sequence 2, Appli
7	1293.5	66.9	354	US-10-790-589-2	Sequence 2, Appli
8	1293.5	66.9	354	US-10-670-863-1	Sequence 2, Appli
9	613	31.7	360	US-10-408-167A-2	Sequence 1, Appli
10	580	30.0	348	US-10-408-167A-2	Sequence 2, Appli
11	98.5	5.1	391	US-10-282-122A-1164	Sequence 1164, Ap
12	97.5	5.0	697	US-10-467-53A-39	Sequence 69762, A
13	96.5	5.0	661	US-10-289-762-36	Sequence 36, Appli
14	95	4.9	681	US-10-362-226-1	Sequence 1, Appli
15	95	4.9	687	US-09-842-758-24	Sequence 24, Appli

16	95	4.9	687	12	US-10-174-333-24	Sequence 24, Appli
17	95	4.9	704	10	US-09-842-758-26	Sequence 26, Appli
18	95	4.9	704	12	US-10-174-333-26	Sequence 26, Appli
19	95	4.9	704	12	US-09-972-211-98	Sequence 98, Appli
20	95	4.9	704	12	US-09-833-782-2	Sequence 2, Appli
21	95	4.9	704	12	US-10-096-625-98	Sequence 98, Appli
22	95	4.9	704	14	US-10-192-207-1	Sequence 1, Appli
23	95	4.9	704	14	US-10-192-445-1	Sequence 1, Appli
24	95	4.9	714	16	US-10-311-035-4	Sequence 4, Appli
25	95	4.9	884	12	US-10-424-599-214353	Sequence 214353,
26	94	4.9	537	16	US-10-322-281-368	Sequence 368, App
27	93	4.8	525	12	US-10-282-122A-78916	Sequence 78916, A
28	92.5	4.8	272	16	US-10-437-963-152834	Sequence 152834,
29	92.5	4.8	638	12	US-10-437-963-152834	Sequence 22, Appli
30	92.5	4.8	638	12	US-10-174-333-22	Sequence 22, Appli
31	92	4.8	645	16	US-10-322-281-364	Sequence 364, App
32	91	4.7	653	16	US-10-437-963-173197	Sequence 173197,
33	91	4.7	1063	15	US-10-161-493-64	Sequence 64, Appli
34	91	4.7	1063	15	US-10-161-493-66	Sequence 66, Appli
35	90.5	4.7	384	12	US-10-206-576-276	Sequence 276, App
36	90.5	4.7	430	12	US-10-206-576-274	Sequence 274, App
37	90.5	4.7	695	15	US-10-312-273-91	Sequence 91, Appli
38	90	4.7	418	12	US-10-424-599-189389	Sequence 189389,
39	89.5	4.6	476	16	US-10-661-809-11	Sequence 11, Appli
40	89	4.6	527	9	US-09-738-626-4803	Sequence 4803, Ap
41	89	4.6	607	15	US-10-108-260A-3710	Sequence 3710, Ap
42	88.5	4.6	624	12	US-09-972-211-24	Sequence 24, Appli
43	88.5	4.6	624	12	US-10-096-625-24	Sequence 24, Appli
44	88.5	4.6	704	12	US-09-972-211-12	Sequence 12, Appli
45	88.5	4.6	704	12	US-09-972-211-14	Sequence 14, Appli

## ALIGNMENTS

RESULT 1  
US-09-949-434-4  
; Sequence 4, Application US/09949434  
; Patent No. US20020028495A1  
; GENERAL INFORMATION:  
; APPLICANT: Eastman, Alan R.  
; APPLICANT: Krieser, Ronald J.  
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS  
; FILE REFERENCE: DC-0137  
; CURRENT APPLICATION NUMBER: US/09/949,434  
; PRIOR FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: 09/574,942  
; PRIOR FILING DATE: 2000-05-19  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 357  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-949-434-4

Query Match	100.0%	Score	1933	DB	9	Length	357
Best Local Similarity	100.0%	Pred. No.	1.3e-186				
Matches	357	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
QY	1	MMARLRTSFALLFLGLFGVLCGAATISCRNEGKAVDWFTFYKLPKRNKESGETGLEYL	60				
Db	1	MMARLRTSFALLFLGLFGVLCGAATISCRNEGKAVDWFTFYKLPKRNKESGETGLEYL	60				
QY	61	YLDSTTRSRKSEQLMNDTKSVLGRTLQQLYAYASKNNNTAYLYNDGVKPVNYSRY	120				
Db	61	YLDSTTRSRKSEQLMNDTKSVLGRTLQQLYAYASKNNNTAYLYNDGVKPVNYSRY	120				
QY	121	GHTKGLLLNNRVQGFLLHSIPQFPPIPEEGYDYPFTGRNGQSGICITFKNQVEAIDS	180				
Db	121	GHTKGLLLNNRVQGFLLHSIPQFPPIPEEGYDYPFTGRNGQSGICITFKNQVEAIDS	180				
QY	181	QLLVNPNVYSCSIPATPHQELIHPQLCTRASSSEIPGRLLITLQSAQGGKFLHFAKSD	240				

Db 181 QLLVCNPNVYSCSIPATFHQELHMPOLCTRASSSEIPGRLLTTLSAQGQKFLHFAKSD 240  
QY 241 SFLLDDIFAAWMAQRLKTHLLTETWQRKQELPNSCLSPYHYVNIKAIKLSRHSYFSSYQD 300  
Db 241 SFLLDDIFAAWMAQRLKTHLLTETWQRKQELPNSCLSPYHYVNIKAIKLSRHSYFSSYQD 300  
QY 301 HAKWCISQKGTNRWTCIGDLNRSPhoAfrSGGFICTQNWQIYQAFQGLVLYESCK 357  
Db 301 HAKWCISQKGTNRWTCIGDLNRSPhoAfrSGGFICTQNWQIYQAFQGLVLYESCK 357

RESULT 2  
US-10-790-709-4  
; Sequence 4, Application US/10240709  
; Publication No. US20030212023A1  
; GENERAL INFORMATION:  
; APPLICANT: Eastman, Alan R.  
; APPLICANT: Krieser, Ronald J.  
; APPLICANT: Trustees of Dartmouth College  
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND cDNAs  
; FILE REFERENCE: DC-0154  
; CURRENT APPLICATION NUMBER: US/10/240,709  
; CURRENT FILING DATE: 2002-10-01  
; PRIOR APPLICATION NUMBER: 09/541,840  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: 09/574,942  
; PRIOR FILING DATE: 2000-05-19  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 357  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-790-709-4

Query Match 100.0%; Score 1933; DB 12; Length 357;  
Best Local Similarity 100.0%; Pred. No. 1.3e-186;  
Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMARLLRTSFALLFLGLFGVLGAATISCRNEEGKAVDWFTFYKLPKRQKESGETGLEYL 60  
Db 1 MMARLLRTSFALLFLGLFGVLGAATISCRNEEGKAVDWFTFYKLPKRQKESGETGLEYL 60

QY 61 YLDSTTRSWRKSEQLMNDTKSVLGRTLQQLYEAYASKNNNTAYLIYNDGVKPKVNYSRKY 120  
Db 61 YLDSTTRSWRKSEQLMNDTKSVLGRTLQQLYEAYASKNNNTAYLIYNDGVKPKVNYSRKY 120

QY 121 GHTKGLLLNRRVQGFLLHSIPQFPPIPEEGYDPTGRRNGQSGICITFKYNOYEADSD 180  
Db 121 GHTKGLLLNRRVQGFLLHSIPQFPPIPEEGYDPTGRRNGQSGICITFKYNOYEADSD 180

QY 181 QLLVCNPNVYSCSIPATFHQELHMPOLCTRASSSEIPGRLLTTLSAQGQKFLHFAKSD 240  
Db 181 QLLVCNPNVYSCSIPATFHQELHMPOLCTRASSSEIPGRLLTTLSAQGQKFLHFAKSD 240

QY 241 SFLLDDIFAAWMAQRLKTHLLTETWQRKQELPNSCLSPYHYVNIKAIKLSRHSYFSSYQD 300  
Db 241 SFLLDDIFAAWMAQRLKTHLLTETWQRKQELPNSCLSPYHYVNIKAIKLSRHSYFSSYQD 300

QY 301 HAKWCISQKGTNRWTCIGDLNRSPhoAfrSGGFICTQNWQIYQAFQGLVLYESCK 357  
Db 301 HAKWCISQKGTNRWTCIGDLNRSPhoAfrSGGFICTQNWQIYQAFQGLVLYESCK 357

RESULT 3  
US-10-790-589-4  
; Sequence 4, Application US/10790589  
; Publication No. US20040142376A1  
; GENERAL INFORMATION:  
; APPLICANT: Eastman, Alan R.  
; APPLICANT: Krieser, Ronald J.  
; APPLICANT: Trustees of Dartmouth College

; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND cDNAs  
; FILE REFERENCE: DC-0154  
; CURRENT APPLICATION NUMBER: US/10/790,589  
; CURRENT FILING DATE: 2004-03-01  
; PRIOR APPLICATION NUMBER: 09/541,840  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: 09/574,942  
; PRIOR FILING DATE: 2000-05-19  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 357  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-790-589-4

Query Match 100.0%; Score 1933; DB 16; Length 357;  
Best Local Similarity 100.0%; Pred. No. 1.3e-186;  
Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMARLLRTSFALLFLGLFGVLGAATISCRNEEGKAVDWFTFYKLPKRQKESGETGLEYL 60  
Db 1 MMARLLRTSFALLFLGLFGVLGAATISCRNEEGKAVDWFTFYKLPKRQKESGETGLEYL 60

QY 61 YLDSTTRSWRKSEQLMNDTKSVLGRTLQQLYEAYASKNNNTAYLIYNDGVKPKVNYSRKY 120  
Db 61 YLDSTTRSWRKSEQLMNDTKSVLGRTLQQLYEAYASKNNNTAYLIYNDGVKPKVNYSRKY 120

QY 121 GHTKGLLLNRRVQGFLLHSIPQFPPIPEEGYDPTGRRNGQSGICITFKYNOYEADSD 180  
Db 121 GHTKGLLLNRRVQGFLLHSIPQFPPIPEEGYDPTGRRNGQSGICITFKYNOYEADSD 180

QY 181 QLLVCNPNVYSCSIPATFHQELHMPOLCTRASSSEIPGRLLTTLSAQGQKFLHFAKSD 240  
Db 181 QLLVCNPNVYSCSIPATFHQELHMPOLCTRASSSEIPGRLLTTLSAQGQKFLHFAKSD 240

QY 241 SFLLDDIFAAWMAQRLKTHLLTETWQRKQELPNSCLSPYHYVNIKAIKLSRHSYFSSYQD 300  
Db 241 SFLLDDIFAAWMAQRLKTHLLTETWQRKQELPNSCLSPYHYVNIKAIKLSRHSYFSSYQD 300

QY 301 HAKWCISQKGTNRWTCIGDLNRSPhoAfrSGGFICTQNWQIYQAFQGLVLYESCK 357  
Db 301 HAKWCISQKGTNRWTCIGDLNRSPhoAfrSGGFICTQNWQIYQAFQGLVLYESCK 357

RESULT 4  
US-10-670-863-3  
; Sequence 3, Application US/10670863  
; Publication No. US20040157239A1  
; GENERAL INFORMATION:  
; APPLICANT: Tanuma, Sei-ichi  
; APPLICANT: Shiohawa, Daisuke  
; TITLE OF INVENTION: Novel Deoxyribonuclease, Gene Encoding Thereof and Use Thereof  
; FILE REFERENCE: 224589  
; CURRENT APPLICATION NUMBER: US/10/670,863  
; CURRENT FILING DATE: 2003-09-25  
; PRIOR APPLICATION NUMBER: JP 11-230870  
; PRIOR FILING DATE: 1999-08-17  
; PRIOR APPLICATION NUMBER: US 09/807,784  
; PRIOR FILING DATE: 2001-04-17  
; NUMBER OF SEQ ID NOS: 15  
; SEQ ID NO 3  
; LENGTH: 361  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-670-863-3

Query Match 99.6%; Score 1926; DB 16; Length 361;  
Best Local Similarity 99.7%; Pred. No. 6.5e-186;  
Matches 356; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MMARLLRTSFALLFLGLFGVLGAATISCRNEEGKAVDWFTFYKLPKRQKESGETGLEYL 60

```

Db      5  MMARLLRTSFALLFLGLFGLVGAATISCRNEBEGKAVDWFYFKLPKRONKESGETGLEYL 64
Qy      61  YLDSTTRSWRKSEQLMNDTKSVLGRITLQOLYEAYASKNNTAYLIYNDGVPKPNYSRKY 120
Db      65  YLDSTTRSWRKSEQLMNDTKSVLGRITLQOLYEAYASKNNTAYLIYNDGVPKPNYSRKY 124
Qy      121  GHTKGLLLNWRVQGWFLIHSIPQPPPIPEEGYDPTGRRNGSGICITFKKNOYEADIS 180
Db      125  GHTKGLLLNWRVQGWFLIHSIPQPPPIPEEGYDPTGRRNGSGICITFKKNOYEADIS 184
Qy      181  QLLVCPNVYSCIPATFHQELHMPOLCTRASSEIPGRLLTTLQSAQOGKFLHFAKSD 240
Db      185  QLLVCPNVYSCIPATFHQELHMPOLCTRASSEIPGRLLTTLQSAQOGKFLHFAKSD 244
Qy      241  SFLLDDIFAANMAORLKHLLTETWQRKQELPNSCLPYHVYNIKAIKLSRHSYFSSYQD 300
Db      245  SFLLDDIFAANMAORLKHLLTETWQRKQELPNSCLPYHVYNIKAIKLSRHSYFSSYQD 304
Qy      301  HAKWCISQKGTNRWTCIGDLNRSPHOAFRSGGFICTQNWQIYQAFQGLVLYYESCK 357
Db      305  HAKWCISQKGTNRWTCIGDLNRSPHOAFRSGGFICTQNWQIYQAFQGLVLYYESCK 361

```

RESULT 5

```

US-09-949-434-2
; Sequence 2, Application US/09949434
; Patent No. US2002028495A1
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Krieser, Ronald J.
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCE: DC-0137
; CURRENT APPLICATION NUMBER: US/09/949,434
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-949-434-2

```

Query Match 66.9%; Score 1293.5; DB 9; Length 354;  
 Best Local Similarity 66.4%; Pred. No. 6.7e-122;  
 Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

```

Qy      1  MMARLLRTSFALLFLGLFGLVGAATISCRNEBEGKAVDWFYFKLPKRONKESGETGLEYL 60
Db      1  MTAKPLRTVLSLFFALLSGVLGTPETSCRNEYGEAVDWFYFKLPKRTSKASEAGLQYL 60
Qy      61  YLDSTTRSWRKSEQLMNDTKSVLGRITLQOLYEAYASKNNTAYLIYNDGVPKPNYSRKY 120
Db      61  YLDSTRTQWNKSLYLINSTRSALGRITLQOLYDTHNS--TNDTAYLIYNDGVPKPNYSRQY 119
Qy      121  GHTKGLLLNWRVQGWFLIHSIPQPPPIPEEGYDPTGRRNGSGICITFKKNOYEADIS 180
Db      120  GHAKGLLVNRTQGWFLIHSVPKPPV--HGYYPTSGRRYGTGICITFGYSQFBEIDF 177
Qy      181  QLLVCPNVYSCIPATFHQELHMPOLCTRASSEIPGRLLTTLQSAQOGKFLHFAKSD 240
Db      178  QLLVLPQNIYSCIPSTFHWKLLYMPKNCANSSSLKIPVRYLAELHSAQGLNFVHFAKSS 237
Qy      241  SFLLDDIFAANMAORLKHLLTETWQRKQELPNSCLPYHVYNIKAIKLSRHSYFSSYQD 300
Db      238  FYDDIFTGIAQKLTLLAQTWQKKQELPNSCLPYHVYNIKSIQVTSKYSFSSRQD 297
Qy      301  HAKWCISQKGTNRWTCIGDLNRSPHOAFRSGGFICTQNWQIYQAFQGLVLYYESCK 357
Db      298  HSKWCVSIKGSANRWTCIGDLNRSLHQALRGGGFICTKNHYYQAFHKLRYGFCCK 354

```

RESULT 7

```

US-10-790-589-2
; Sequence 2, Application US/10790589
; Publication No. US20040142376A1
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Krieser, Ronald J.
; APPLICANT: Trustees of Dartmouth College
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCE: DC-0154
; CURRENT APPLICATION NUMBER: US/10/790,589
; PRIOR FILING DATE: 2004-03-01
; PRIOR APPLICATION NUMBER: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 354

```

RESULT 6

```

US-10-240-709-2
; Sequence 2, Application US/10240709
; Publication No. US20030212023A1
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Krieser, Ronald J.
; APPLICANT: Trustees of Dartmouth College
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCE: DC-0154
; CURRENT APPLICATION NUMBER: US/10/240,709
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-240-709-2

```

Query Match 66.9%; Score 1293.5; DB 12; Length 354;  
 Best Local Similarity 66.4%; Pred. No. 6.7e-122;  
 Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

```

Qy      1  MMARLLRTSFALLFLGLFGLVGAATISCRNEBEGKAVDWFYFKLPKRONKESGETGLEYL 60
Db      1  MTAKPLRTVLSLFFALLSGVLGTPETSCRNEYGEAVDWFYFKLPKRTSKASEAGLQYL 60
Qy      61  YLDSTTRSWRKSEQLMNDTKSVLGRITLQOLYEAYASKNNTAYLIYNDGVPKPNYSRKY 120
Db      61  YLDSTRTQWNKSLYLINSTRSALGRITLQOLYDTHNS--TNDTAYLIYNDGVPKPNYSRQY 119
Qy      121  GHTKGLLLNWRVQGWFLIHSIPQPPPIPEEGYDPTGRRNGSGICITFKKNOYEADIS 180
Db      120  GHAKGLLVNRTQGWFLIHSVPKPPV--HGYYPTSGRRYGTGICITFGYSQFBEIDF 177
Qy      181  QLLVCPNVYSCIPATFHQELHMPOLCTRASSEIPGRLLTTLQSAQOGKFLHFAKSD 240
Db      178  QLLVLPQNIYSCIPSTFHWKLLYMPKNCANSSSLKIPVRYLAELHSAQGLNFVHFAKSS 237
Qy      241  SFLLDDIFAANMAORLKHLLTETWQRKQELPNSCLPYHVYNIKAIKLSRHSYFSSYQD 300
Db      238  FYDDIFTGIAQKLTLLAQTWQKKQELPNSCLPYHVYNIKSIQVTSKYSFSSRQD 297
Qy      301  HAKWCISQKGTNRWTCIGDLNRSPHOAFRSGGFICTQNWQIYQAFQGLVLYYESCK 357
Db      298  HSKWCVSIKGSANRWTCIGDLNRSLHQALRGGGFICTKNHYYQAFHKLRYGFCCK 354

```

```

; TYPE: PRT
; ORGANISM: Mus sp.
US-10-790-589-2

Query Match 66.9%; Score 1293.5; DB 16; Length 354;
Best Local Similarity 66.4%; Pred. No. 6.7e-122;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

QY 1 MMARLRTSFALLFLGLFGVLGAATISCRNEGKAVDWFYFKLPKRONKESGETGLEYL 60
Db 1 MTAKPLRIVLSLLFFALSGLVLTPEISCRNEGVEAVDWFYFKLPKRTSKASEAGLQYL 60

QY 61 YLDSTTRSWRKSEQLMNDTKSVLGRITLQOLYEAYASKSNNTAYLIYNDGVKPVNYSRKY 120
Db 61 YLDSTTRQWNKSLYLINSTRSALCRILQHLIDYTHNS-TNDTAYLIYNDGVKPVNYSRQY 119

QY 121 GHTKGLLLNRRVOGFWLHLSIPOFPPIPEGVDYPTGRRNGOSGICITFKINOYEIDS 180
Db 120 GHAKGLLLVNRITQGFLLHLSVPKPPV--HGVEYPTSGRRYGGTGICITFGYSQFEIDF 177

QY 181 QLLVCPNVYSCISIPATFQELIHMQLCTRASSSIPGRLITLQASOGQKFLHFAKSD 240
Db 178 QLLVLOQNIYSCIPSTFTFWKLLIYMPRMCANSSLLIPVRYLAELHSAQGLNPFVHFAKSD 237

QY 241 SFDDIPFAAWAORLKTLLHTEWORKROELPNCNSLPYHVYNIKAIKLSRHSYFSSYQD 300
Db 238 FYTDDIIFGTWIAQKLKTHLLAQTWQKKQELPNCNSLPYHVYNIKISIGVTSKYSFSSROD 297

QY 301 HAKWCISQKGTQNRWTICGLNRSHPHQAQSGFICTQNWQIYQAQGLVLYTESCK 357
Db 298 HSKWCVSIKGSANEWTICGLNRSLSHQAQSGGFICTKNHYIYQAQFKLYLYRGFCK 354

RESULT 8
US-10-670-863-1
; Sequence 1, Application US/10670863
; Publication No. US20040157239A1
; GENERAL INFORMATION:
; APPLICANT: Tanuma, Sei-ichi
; APPLICANT: Shikawa, Daisuke
; TITLE OF INVENTION: Novel Deoxyribonuclease, Gene Encoding Thereof and Use Thereof
; FILE REFERENCE: 224589
; CURRENT APPLICATION NUMBER: US/10/670,863
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: JP 11-230870
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: US 09/807,784
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 15
; SEQ ID NO 1
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-670-863-1

Query Match 66.9%; Score 1293.5; DB 16; Length 354;
Best Local Similarity 66.4%; Pred. No. 6.7e-122;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

QY 1 MMARLRTSFALLFLGLFGVLGAATISCRNEGKAVDWFYFKLPKRONKESGETGLEYL 60
Db 1 MTAKPLRIVLSLLFFALSGLVLTPEISCRNEGVEAVDWFYFKLPKRTSKASEAGLQYL 60

QY 61 YLDSTTRSWRKSEQLMNDTKSVLGRITLQOLYEAYASKSNNTAYLIYNDGVKPVNYSRKY 120
Db 61 YLDSTTRQWNKSLYLINSTRSALCRILQHLIDYTHNS-TNDTAYLIYNDGVKPVNYSRQY 119

QY 121 GHTKGLLLNRRVOGFWLHLSIPOFPPIPEGVDYPTGRRNGOSGICITFKINOYEIDS 180
Db 120 GHAKGLLLVNRITQGFLLHLSVPKPPV--HGVEYPTSGRRYGGTGICITFGYSQFEIDF 177

QY 181 QLLVCPNVYSCISIPATFQELIHMQLCTRASSSIPGRLITLQASOGQKFLHFAKSD 240

```

Db 181 VNYQLEGIFAQEPDLENVKGHHVQBPWNSSITLTSQAGAVFQSFQAKTSKFGDDLYS 240  
 QY 249 AWMAQLKTHLLTETWQRKQELPNSCSLPYHVNKAIKLSRHS--YFSSYQDHAKWCI 306  
 Db 241 GMLAALGTLNLOVQFWHTKVTGILPNSCDIWOVLNVNQIAFPAGPAGPSFNSTEDHSKWCY 300  
 QY 307 SQKGTNRWTCIGDLNRSFQAFRSGGFICTQNWQIQAQFGLVLYYESC 356  
 Db 301 SPKGP---WTCVGMNRNQEEQGGTLCALPALWKAQFPLVKNYQPC 347

## RESULT 10

US-10-408-765A-1164  
 ; Sequence 1164, Application US/10408765A  
 ; Publication No. US20040101874A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ghosh, Soumitra S.  
 ; APPLICANT: Faby, Eoin D.  
 ; APPLICANT: Zhang, Bing  
 ; APPLICANT: Gibson, Bradford W.  
 ; APPLICANT: Taylor, Steven W.  
 ; APPLICANT: Glenn, Gary M.  
 ; APPLICANT: Warnock, Dale E.  
 ; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
 ; FILE REFERENCE: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME  
 ; CURRENT APPLICATION NUMBER: US/10/408,765A  
 ; CURRENT FILING DATE: 2003-04-04  
 ; NUMBER OF SEQ ID NOS: 3077  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1164  
 ; LENGTH: 348  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-408-765A-1164

Query Match 30.0%; Score 580; DB 16; Length 348;  
 Best Local Similarity 36.3%; Pred. No. 1.1e-49;  
 Matches 127; Conservative 56; Mismatches 143; Indels 24; Gaps 7;

QY 13 LFLGLFVLGAATISCRNEEGKAVDWFYFKLPK-RQNKESGETGLFYLYLDSTTRSWRK 71  
 Db 4 LLLAALLCVPAAGALTCYDGSQCPVDWFWVKLPALRGSGEAAQKGLQKYLDESSGGWRD 63  
 QY 72 SEQLMNDTKSVLGRTLQOLYEAVASKNNNTAYLIYNDGVKPP--VNYSRKYGHTKGLLLW 129  
 Db 64 GRALINSPEGAVERSLQPLTR---SNTSLAFLIYNDQPPQPSKAQDSSMRGHTKGVL 120  
 QY 130 NRVOGFWLIHSIPQF-PPIPEEGYDYPPTGRRNGQSGICITFKYNOYEAIDSOLLVCNPN 188  
 Db 121 DHDGGFWLVHVSFVNPFPASSAAYSWPHSACTYGTLLC-----KQLTYPWP 168  
 QY 189 VYSCSPATPHOLHIMPQLCTRASSEIPGRLLTLOSAGQKFLHFAKSDSFLDDIFA 248  
 Db 169 VNYQLEGIFAQEPDLENVKGHHVQBPWNSSITLTSQAGAVFQSFQAKTSKFGDDLYS 228  
 QY 249 AWMAQLKTHLLTETWQRKQELPNSCSLPYHVNKAIKLSRHS--YFSSYQDHAKWCI 306  
 Db 229 GMLAALGTLNLOVQFWHTKVTGILPNSCDIWOVLNVNQIAFPAGPAGPSFNSTEDHSKWCY 288  
 QY 307 SQKGTNRWTCIGDLNRSFQAFRSGGFICTQNWQIQAQFGLVLYYESC 356  
 Db 289 SPKGP---WTCVGMNRNQEEQGGTLCALPALWKAQFPLVKNYQPC 335

## RESULT 11

US-10-282-122A-69762  
 ; Sequence 69762, Application US/10282122A  
 ; Publication No. US20040029129A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Liangsu  
 ; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl  
 ; APPLICANT: Haseibeck, Robert  
 ; APPLICANT: Ohlsen, Kari  
 ; APPLICANT: Zyskind, Judith  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John  
 ; APPLICANT: Carr, Grant  
 ; APPLICANT: Yamamoto, Robert  
 ; APPLICANT: Forsyth, R.  
 ; APPLICANT: Xu, H.  
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 ; FILE REFERENCE: ELITRA 034A  
 ; CURRENT APPLICATION NUMBER: US/10/282,122A  
 ; CURRENT FILING DATE: 2003-02-20  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/230,335  
 ; PRIOR FILING DATE: 2000-09-06  
 ; PRIOR APPLICATION NUMBER: 60/230,347  
 ; PRIOR FILING DATE: 2000-09-09  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/267,636  
 ; PRIOR FILING DATE: 2001-02-09  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 78614  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 69762  
 ; LENGTH: 391  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas syringae  
 US-10-282-122A-69762

Query Match 5.1%; Score 98.5; DB 12; Length 391;  
 Best Local Similarity 24.0%; Pred. No. 0.71;  
 Matches 63; Conservative 33; Mismatches 110; Indels 57; Gaps 11;

QY 15 LGLFVLGAATISCRNEEGKAVDWFYFKLPKRNKESGET-----GLEYLYL-DS 64  
 Db 10 LGL-GANGAATVQLAKAGVDVIGVDYAPPHQTQSSHGDTIRLTVSGEGPQLPLVRS 68  
 QY 65 TTRSWRKSEQLMNDTK---SVLGRTLQOLYEAVASKNNNTAYLIYNDGVKPPVNSRKY 120  
 Db 69 SHRIWELEALSGESLFEQCGVLVMTSSPSYDAADA-----DDFTHKTIALARAY 118  
 QY 121 GHTKGLLLNWRVQFWLIHSIPQFPI-----PEEGYDYPPTGRRNGQSGICITFK 171  
 Db 119 GVEHQVIMATAIR-----ERFPQFAPVLDTAIGYFPEGQFVRPER-----CIGVQ 164  
 QY 172 YNQYEAIDSOLL---VCNPNVYSCSIPATFHQELHMPQLCTRAS--SSEIPGRLLTTL 225  
 Db 165 LRLAAGHARLTHETVTHLQYGEQVRIITRKSGSIADKVVVSVAGMWSTELLGAPFSDL 224  
 QY 226 QSAQOQKFLHFAKSDSFLDDIFA 248  
 Db 225 LRVCRCQKLFWRERQ---DAVEA 244

RESULT 12  
 US-10-467-534-39  
 ; Sequence 39, Application US/10467534  
 ; Publication No. US20040131625A1  
 ; GENERAL INFORMATION:

```

; APPLICANT: Berthet, Francois-Xavier Jacques
; APPLICANT: Lobet, Yves
; APPLICANT: Poolman, Jan
; APPLICANT: Veilant, Vincent Georges Christian Louis
; TITLE OF INVENTION: Vaccine Composition
; FILE REFERENCE: B45261
; CURRENT APPLICATION NUMBER: US/10/467,534
; CURRENT FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103169.9
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 697
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-467-534-39

Query Match      5.0%; Score 97.5; DB 16; Length 697;
Best Local Similarity 24.9%; Pred. No. 2.1;
Matches 62; Conservative 25; Mismatches 109; Indels 53; Gaps 12;

QY 114 VNSRK-----YGHTKGLLWN-RVQGF-WLIH-----SIQFPPIPEGVDYPTG 158
DB 85 VNYRKILICDYEYEDTSCLLNGKCSLPWFIGSTTISPSIIHKGYISTSEG 144
QY 159 RRGQSGICITPKYQYBAIDSQLVCNPNVYSCSIPATFHOEL-----IHPOLCTRA 212
DB 145 ---PQHICLSGDLKYSS-DSVLSMGPSRLSICNTFVLLLPQISIMEIPEKPIITFRG 200
QY 213 SSSEIPGRLLTTLQSAQOKFLHFAKSDSFLDDIFAANWAQELKTHLLTETWQRKQELP 272
DB 201 GSGFLGSLVGSYSPISKK---HCSITLFLDGFHKGIG-----LGNMRFSSQNP 250
QY 273 SNCSLPYHYNKAIKLSRHSYFSSYODHAKWCISOKGT-----KRWTCIGDL----- 332
DB 251 SN-----AINIK-----SYAH-RLAIDSSGAKRYRLHGDFTFSKRAHLAG 292
QY 333 GFICTQNWQ 341
DB 293 EPHLSDSWE 301

RESULT 13
US-10-289-762-36
; Sequence 36, Application US/10289762
; Publication No. US20040006218A1
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/10/289,762
; CURRENT FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 36
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-289-762-36

Query Match      5.0%; Score 96.5; DB 15; Length 661;
Best Local Similarity 19.7%; Pred. No. 2.5;
Matches 76; Conservative 59; Mismatches 157; Indels 93; Gaps 19;

QY 4 RLRTSFALLFLGFLVGLGATISCRNEEG--KAVDMFTFKLPKQKNSGEGTGLE--Y 59
DB 9 RLMKRCF--LFLASFLVLMGSSADALHQAQKKNVLSHFK-----SVSGIVTIEDGV 60
QY 60 LYLDSTTRSWRKSEQLMNDTKSVLGRTLQQLYEA-----YASKSNNTAYLIYNDGVPKV 114

; APPLICANT: Berthet, Francois-Xavier Jacques
; APPLICANT: Lobet, Yves
; APPLICANT: Poolman, Jan
; APPLICANT: Veilant, Vincent Georges Christian Louis
; TITLE OF INVENTION: Vaccine Composition
; FILE REFERENCE: B45261
; CURRENT APPLICATION NUMBER: US/10/467,534
; CURRENT FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103169.9
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 697
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-467-534-39

Query Match      5.0%; Score 97.5; DB 16; Length 697;
Best Local Similarity 24.9%; Pred. No. 2.1;
Matches 62; Conservative 25; Mismatches 109; Indels 53; Gaps 12;

QY 114 VNSRK-----YGHTKGLLWN-RVQGF-WLIH-----SIQFPPIPEGVDYPTG 158
DB 85 VNYRKILICDYEYEDTSCLLNGKCSLPWFIGSTTISPSIIHKGYISTSEG 144
QY 159 RRGQSGICITPKYQYBAIDSQLVCNPNVYSCSIPATFHOEL-----IHPOLCTRA 212
DB 145 ---PQHICLSGDLKYSS-DSVLSMGPSRLSICNTFVLLLPQISIMEIPEKPIITFRG 200
QY 213 SSSEIPGRLLTTLQSAQOKFLHFAKSDSFLDDIFAANWAQELKTHLLTETWQRKQELP 272
DB 201 GSGFLGSLVGSYSPISKK---HCSITLFLDGFHKGIG-----LGNMRFSSQNP 250
QY 273 SNCSLPYHYNKAIKLSRHSYFSSYODHAKWCISOKGT-----KRWTCIGDL----- 332
DB 251 SN-----AINIK-----SYAH-RLAIDSSGAKRYRLHGDFTFSKRAHLAG 292
QY 333 GFICTQNWQ 341
DB 293 EPHLSDSWE 301

RESULT 14
US-10-362-226-1
; Sequence 1, Application US/10362226
; Publication No. US20040014939A1
; GENERAL INFORMATION:
; APPLICANT: DAICHI PHARMACEUTICAL CO., LTD
; TITLE OF INVENTION: No. US20040014939A1el1 protease gene
; FILE REFERENCE: GP01-1023
; CURRENT APPLICATION NUMBER: US/10/362,226
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: JP P2000-252503
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 681
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-362-226-1

Query Match      4.9%; Score 95; DB 15; Length 681;
Best Local Similarity 18.7%; Pred. No. 3.7;
Matches 61; Conservative 49; Mismatches 117; Indels 100; Gaps 14;

QY 20 VLGAATISCRNE---EGKAVDWFTFYKLPKQKKE---SGETGLEYLYLDSTTRSWRKS 72
DB 314 ILNLKKCKCKRGFEYDGGKINAWDLYYYMTQTEELKYSIDQBFLKEFFIEVVT----- 367
QY 73 EOLMNDTKSVLGRTLQQLYEAASKSNNTAYLI-----YNDGVPK--PVNYSR 118
DB 368 EGLNTYQELLGLSFEQMTDAHWNKSVTLVYVKDKATGEVLGQFLDLYPREGKYNHAA 427
QY 119 KYCHTKGLLLMNRVQGVFLIHSIPOPPPIPEGVDYVPTGRRNGQSGICITFKYNOYEA 178
DB 428 CFGLQPGCLL-----PDGSRMVAALVNF----- 453
QY 179 DSQLLVCNPNVYSCSIPATFHOELIH-MPOLCTRASS-----EIPGRLLTT- 224
DB 454 -SQVAGRESLLRHDEVRTYFHFHGVHMOICQATDFARFSGTNVETDFVFPQMLENW 512
QY 225 LQSAQOKFL--HFAKSDSFLDDIFAANWAQRI--KTHLLT---ETWQRKQELPNSC 278
DB 513 VMDVDSLRSLKXHYKDGSPADDDLLKLVASRLVNTGLTLRLQIVLSKVDQSLHTNTSL- 571
QY 279 YHYNYNKAIKLSRHSYFSSYQDHAKWC 305
DB 572 -----DAASEYAKYC 581

```





***This Page Blank (uspto)***

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2004, 00:08:06 / Search time 29,1224 Seconds  
(without alignments)  
1179.175 Million cell updates/sec

Title: US-10-790-589-4

Perfect score: 1933

Sequence: 1 MWARLLRTSFALLFLGLFGV.....QNWQIYQAFQGLVLYESCK 357

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 78.\*

2: PIR1.\*

3: PIR2.\*

4: PIR3.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1297.5	67.1	356	JC7131	deoxyribonuclease
2	613	31.7	360	J50206	deoxyribonuclease
3	593	30.7	364	J50205	deoxyribonuclease
4	580	30.0	348	T45071	hypothetical prote
5	490.5	25.4	375	T19038	hypothetical prote
6	413	21.4	516	S40996	hypothetical prote
7	319	16.5	238	S44793	F09G8.2 protein -
8	117	6.1	344	A44164	secreted glycoprot
9	97.5	5.0	697	H71525	probable outer mem
10	96.5	5.0	704	C81542	conserved hypothet
11	95	4.9	970	S63059	hypothetical prote
12	94.5	4.9	2295	B71621	probable membrane
13	94	4.9	523	DEFFG6	Glucose-6-phosphat
14	93.5	4.8	394	E81286	probable polysacch
15	93	4.8	526	F82873	hypothetical prote
16	93	4.8	6689	S55024	nebulin, skeletal
17	92	4.8	524	A47740	Glucose-6-phosphat
18	92	4.8	641	F90832	terminase large su
19	92	4.8	641	C90901	probable termase
20	92	4.8	641	H85689	terminase large su
21	92	4.8	709	C64057	IgA-specific metal
22	91.5	4.7	305	S73645	probable lipoprote
23	90.5	4.7	695	B72129	probable outer mem
24	90.5	4.7	695	F86493	probable outer mem
25	89.5	4.6	533	D86756	prophage p12 prote
26	89	4.6	831	T05771	beta-galactosidase
27	88.5	4.6	585	T22649	hypothetical prote
28	88.5	4.6	1456	A36563	mannose receptor p
29	88	4.6	424	1 MWV2K3	K3 protein - vacci

30 88 4.6 641 1 JVBPAL DNA-packaging prot  
31 87.5 4.5 700 2 B81682 conserved hypothet  
32 87.5 4.5 877 2 T03098 p97 protein - Toxo  
33 87 4.5 685 1 A47102 system b(0,+) amin  
34 87 4.5 777 2 T25761 hypothetical prote  
35 86 4.4 524 2 S76140 hypothetical prote  
36 86 4.4 704 2 A45985 microsomal endopep  
37 85.5 4.4 508 2 S74537 anthranilate synth  
38 85.5 4.4 521 2 T11166 CDPdiacylglycerol-  
39 85.5 4.4 587 2 T07634 pollen-specific pr  
40 85.5 4.4 951 1 SYECVT valine-tRNA ligase  
41 85.5 4.4 951 2 C91283 valine tRNA synth  
42 85.5 4.4 951 2 E86124 valine tRNA synth  
43 85.5 4.4 951 2 AC1061 valine-tRNA ligase  
44 85.5 4.4 954 2 G64121 valine-tRNA ligase  
45 85.5 4.4 1307 2 AH0446 probable exported

#### ALIGNMENTS

##### RESULT 1

JC7131

deoxyribonuclease I (EC 3.1.21.1) - rat (strain Wistar)

N:Alternate names: deoxyribonuclease II-like acid deoxyribonuclease

C:Species: Rattus norvegicus (Norway rat)

C>Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 09-Dec-2002

C/Accession: JC7131

R:Tanuma, S.; Shiohawa, D.

Biochem. Biophys. Res. Commun. 265, 395-399, 1999

A:Title: Cloning of a cDNA encoding a rat DNase II-like acid DNase.

A:Reference number: JC7131; MUID:20025354; PMID:10558878

A:Accession: JC7131

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-356 <TAN>

A:Cross-references: GB:AF178974; NID:g6470130; PID:g6470131

C:Superfamily: deoxyribonuclease II

C:Keywords: hydrolase

##### Query Match

Best Local Similarity 67.1%; Score 1297.5; DB 2; Length 356;

Matches 236; Conservative 66.3%; Pred. No. 2.5e-103; Indels 1; Gaps 1;

Qy	1	MWARRLLRTSFALLFLGLFGVGLGATISCRNEEGKAVDWFIFYKLPRKQNKESGETGLEYL	60
Db	1	MTAQPLKAALPLFLFVALSGVLGTPVISCINEDGKAVDWFIFYKLPRRTSGGTGMLDYL	60
Qy	61	YLDSTTRESWRKSKQLMNDTKSVLGRITLQLYEAYASKNNNTAVILYNDGVPKPNYSRKY	120
Db	61	YLDSTTRESWRKSKQLMNDTKSVLGRITLQLYEAYASKNNNTAVILYNDGVPKPNYSRKY	120
Qy	121	GHTKGLLLNRRVQGFVLIHSIPQPIPERGYDPPPTGRRNGSGGICITPKYNOYEAIDS	180
Db	120	GKAGLLVNNRVQGFVLIHSIPKFPVPEKGYEYPSGSGRYAQSGLCITIKYSFTIDS	179
Qy	181	QLLVCPNIVSCIPATFFHQLHMPOLCTRASSETPGRLITLTLQSAQOKFTHFAKSD	240
Db	180	QLLVFQPNIVSCIPATFFHQLHMPOLCTRASSETPGRLITLTLQSAQOKFTHFAKST	239
Qy	241	SELDIDFAAMWAQRLKTHLTETWQRKQRLPNCSLPHYVYIKAIKLSRHSYFSYQD	300
Db	240	FYTDDIFAAWIAOKLVHLLVESWQRKNHLPNCSLPHYVYIKAIKLSRHSYFSYQD	299
Qy	301	HAKWCISQKTKNRWTCIGDLNRSFHOAPSGGFCITQNWQIYQAFQGLVLYESCK	356
Db	300	HSKWCYSTKDSQARWTCIGDLNRSFHOAPSGGFCITQNWQIYQAFQGLVLYESCK	355

##### RESULT 2

J50206

deoxyribonuclease II (EC 3.1.22.1) - human

C:Species: Homo sapiens (man)

C;Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change 09-Dec-2002  
C;Accession: J02006  
R;Shiokawa, D.; Tanuma, S.  
Biochem. Biophys. Res. Commun. 247, 864-869, 1998  
A;Title: Cloning of cDNAs encoding porcine and human DNase II.  
A;Reference number: J02005; MUID:98321218; PMID:9647784  
A;Accession: J02006  
A;Molecule type: mRNA  
A;Residues: 1-364 <SHI>  
A;Cross-references: GB:AF060222; NID:g3309154; PIDN:AA39852.1; PID:g3309155  
C;Comment: This enzyme catalyzes the hydrolysis of DNA into 3'-phosphoryl oligonucleotides  
C;Genetics:  
A;Map position: 19p13.2  
C;Superfamily: deoxyribonuclease II  
C;Keywords: hydrolase

Query Match 31.7%; Score 613; DB 2; Length 360;  
Best Local Similarity 36.9%; Pred. No. 1e-44;  
Matches 129; Conservative 61; Mismatches 148; Indels 12; Gaps 6;

QY 13 LFLGLFGLVGAATISCRNEEGKAVDFTFYKLPK-RONKESGETGLYLYLDTSTRWRK 71  
DB 4 LLLAALLCVPAGALTCTGDSGQPDVDFVYKLPALRGSGEAAQGLQYKYLDESSGGWRD 63  
QY 72 SEQLMNDTKSVLGRITLQOLYEAYASKSNNTAYLIYNDGVKPK--VNSRYKHTKGLLLW 129  
DB 64 GRALINSPGAVGRSLQPLYR---SNTSQAFLLYNDQPPQSKAODSMRGTGKGLLL 120  
QY 130 NRVOGFWLHISIPQF--PIPEEGDYDPTGRRNGQSGICITFKYNQYEAIDSLVNCNP 188  
DB 121 DHGGFWLHVSFNFPPPPASSAAYSWPHSACTYGTLLCVSPFPAPQSKMGKQLTYTPW 180  
QY 189 VYSCSIPATPHOELIHPOLCTRASSSEIPGRLLTTLQSAQGGKFTLHFAKSDSFLDDIFA 248  
DB 181 VVNYQLEGIFPAQEFPPDLENVVKGHVSOEPWNSITLTSQAGAVFQSFKFKGDDLYS 240  
QY 249 AWAQRLKTHLLTETWQRKQELPNSCLPYHYVNIKAIKLSRHS--YFSSYQDHAKWCI 306  
DB 241 GMLAALGTNLQVFWHKTGVLPSNCSDIQWLVNQNQIAFPFGPAGPSFNSDEHSKWCV 300  
QY 307 SQGTGKNRWTICGLNRSPhQAFRSGGFICTQNWQIYQAFQGLVLYYESC 356  
DB 301 SPKGP---WTCVGMNMQEORGGGTLCALPALWKAFLVKNYQPC 347

RESULT 3  
J02005  
deoxyribonuclease II (EC 3.1.22.1) - pig  
C;Species: Sus scrofa domestica (domestic pig)  
C;Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change 09-Dec-2002  
C;Accession: J02005  
R;Shiokawa, D.; Tanuma, S.  
Biochem. Biophys. Res. Commun. 247, 864-869, 1998  
A;Title: Cloning of cDNAs encoding porcine and human DNase II.  
A;Reference number: J02005; MUID:98321218; PMID:9647784  
A;Accession: J02005  
A;Molecule type: mRNA  
A;Residues: 1-364 <SHI>  
A;Cross-references: GB:AF060222; NID:g3309152; PIDN:AA39263.1; PID:g3309153  
C;Comment: This enzyme catalyzes the hydrolysis of DNA into 3'-phosphoryl oligonucleotides  
C;Superfamily: deoxyribonuclease II  
C;Keywords: hydrolase

Query Match 30.7%; Score 593; DB 2; Length 364;  
Best Local Similarity 35.2%; Pred. No. 5.4e-43;  
Matches 126; Conservative 66; Mismatches 152; Indels 14; Gaps 7;

QY 6 LRFSFALLFLGLFGLVGAATISCRNEEGKAVDFTFYKLPKRONK-RSGETGLYLYLDS 64  
DB 1 MATLSPLLALLLW-PVGLITCYGDSGQPDVDFVYKLPALHSPGDAQSGLEYKYLDE 59  
QY 65 TTRSWRKSEQLMNDTKSVLGRITLQOLYEAYASKSNNTAYLIYNDGVKPK--PVNSRYKYG 122

DB 60 ESGWRDAGSINGSTGALGRSLPLXLR---NTSQAFLLYNDQPPKYRQSHSNRGGH 115  
QY 123 TKGILLNVRVGVFWLHISIPQF--PIPEEGDYDPTGRRNGQSGICITFKYNQYEAIDSQ 181  
DB 116 TKGVILLDQEGFWLHISVFNPPSSSAAATSWPFSARTYGTGLICVSPFLTQFLNISRQ 175  
QY 182 LLVCNPNVYSCSIPATPHOELIHPOLCTRASSSEIPGRLLTTLQSAQGGKFTLHFAKSDS 241  
DB 176 LTYTPVMDYDKLEGDFAEKFPYLBEEVYKHHVLQEPWNSVTLTSKAGASQSFACKN 235  
QY 242 FLDDIFAAMWAQRLKTHLLTETWQRKQELPNSCLPYHYVNIKAIKLSRHS--YFSSYQ 299  
DB 236 FGDDLYSGWLAELGSLNQVQFQWQSGAGILPNSCSGVQHVLDVTQIAFPGPAGPNFATE 295  
QY 300 DHAKWCIQGTGKNEWTICGLNRSPhQAFRSGGFICTQNWQIYQAFQGLVLYYESCK 357  
DB 296 DSHKWCVA---PERPWTGVMNMRKREHRRGGTLCALPALWKAFLVKNYQPC 350

RESULT 4  
T45071  
hypothetical protein R31240.2 [imported] - human  
C;Species: Homo sapiens (man)  
C;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Dec-2002  
C;Accession: T45071  
R;Lamerdin, J.; McCreedy, P.; Stilwagen, S.; Ramirez, M.; Carrano, A.  
submitted to the EMBL Data Library, November 1996  
A;Description: Characterization by genomic sequence analysis of a gene-rich 111 kb region  
A;Reference number: Z22906  
A;Accession: T45071  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-348 <LAM>  
A;Cross-references: EMBL:AD000092; PIDN:AAB51172.1  
A;Experimental source: cell line 5H12-B; fibroblast  
C;Genetics:  
A;Map position: 19p13.2  
A;Introns: 29/2; 89/3; 116/1; 159/1; 225/1  
C;Superfamily: deoxyribonuclease II

Query Match 30.0%; Score 580; DB 2; Length 348;  
Best Local Similarity 36.3%; Pred. No. 6.6e-42;  
Matches 127; Conservative 56; Mismatches 143; Indels 24; Gaps 7;

QY 13 LFLGLFGLVGAATISCRNEEGKAVDFTFYKLPK-RONKESGETGLYLYLDTSTRWRK 71  
DB 4 LLLAALLCVPAGALTCTGDSGQPDVDFVYKLPALRGSGEAAQGLQYKYLDESSGGWRD 63  
QY 72 SEQLMNDTKSVLGRITLQOLYEAYASKSNNTAYLIYNDGVKPK--VNSRYKHTKGLLLW 129  
DB 64 GRALINSPGAVGRSLQPLYR---SNTSQAFLLYNDQPPQSKAODSMRGTGKGLLL 120  
QY 130 NRVOGFWLHISIPQF--PIPEEGDYDPTGRRNGQSGICITFKYNQYEAIDSLVNCNP 188  
DB 121 DHGGFWLHVSFNFPPPPASSAAYSWPHSACTYGTLLC-----KQLTYTPW 168  
QY 189 VYSCSIPATPHOELIHPOLCTRASSSEIPGRLLTTLQSAQGGKFTLHFAKSDSFLDDIFA 248  
DB 169 VVNYQLEGIFPAQEFPPDLENVVKGHVSOEPWNSITLTSQAGAVFQSFKFKGDDLYS 228  
QY 249 AWAQRLKTHLLTETWQRKQELPNSCLPYHYVNIKAIKLSRHS--YFSSYQDHAKWCI 306  
DB 229 GMLAALGTNLQVFWHKTGVLPSNCSDIQWLVNQNQIAFPFGPAGPSFNSDEHSKWCV 288  
QY 307 SQGTGKNRWTICGLNRSPhQAFRSGGFICTQNWQIYQAFQGLVLYYESC 356  
DB 289 SPKGP---WTCVGMNMQEORGGGTLCALPALWKAFLVKNYQPC 335

RESULT 5  
T19038  
hypothetical protein C07B5.5 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Dec-2002

C:Accession: T19038  
R:Kershaw, J.  
submitted to the EMBL Data Library, October 1994  
A:Reference number: Z19064  
A:Accession: T19038  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-375 <WIL>  
A:Cross-references: EMBL:Z46266; PIDN:CAA86412.1; GSPDB:GN00028; CESP:C07B5.5  
A:Experimental source: clone C07B5  
C:Genetics:  
A:Gene: CESP:C07B5.5  
A:Map position: X  
A:Introns: 35/2; 82/3; 108/3; 266/1; 308/3; 343/3  
C:Superfamily: deoxyribonuclease II

Query Match 25.4%; Score 490.5; DB 2; Length 375;  
Best Local Similarity 32.2%; Pred. No. 3.4e-34;  
Matches 119; Conservative 65; Mismatches 147; Indels 39; Gaps 12;

QY 6 LRTSFALLFLGVLGA-ATISCRNEEGKAVDWTFYKLPKRQKESGET-----GLEY 59  
Db LSPRAVLFL-LLGVSQYAFACKQDSQNDVDWFAVYKMP--IEKDDGSVTGLAGGVAV 59  
QY 60 LYLDTSTR-SWRKSEQLMNDTKSVLGRITLQOLY-----EAVASKSNNTAY 103  
Db 60 YYVDVKKGTLPSTAKTLDNDQALAYTLQYYDKNDKTFHVMYNDPEWGSKSTGK 119  
QY 104 L-----IYNDGVPKPNVGRKYGHTKGLLMNRVQGFLLHISIPQFPPIPEGYDYP 156  
Db 120 LEEILSNRVYNYTHEDDSTAFGHTKGTTFDGTSGVWLHVSPLFPN-PTK-YEYVP 177  
QY 157 TGRNGSGGICITKYNQYEAIDSQLVCNPNVYSCSIPATFHQELIHPQLCTRASSE 216  
Db 178 SGHDYGQMLCMTKYAQLKISGTLFFNRIYSSNPLTNMAADNADLAKAIAGQYQK 237  
QY 217 IPGRLLTTLSAQOGKFLHFAKSDSFLDDIEAAMQRLKTHLTETWQRKQLPSNCS 276  
Db 238 QPFSVLETFNAGYSFTNFAKSEFNADLTIVAPLTKDVLVETW-RRGSEIFLDC 296  
QY 277 LPYHYNKAIKLSRHSYFSSYQDHAKWCISQKGTNRKWTTCIGDINRSPHQAFRSGGFC 336  
Db 297 LTHANDALSITHVSTAFSTYKDHSMHAADMTK-PWVCIGDINRMTSQYVRGGTTC 355  
QY 337 TQN---WQIY 343  
Db 356 ISSSFLWKAY 365

RESULT 6  
S40996  
hypotheical protein K04H4.6 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 03-May-1994 #sequence\_revision 02-Aug-1994 #text\_change 09-Sep-1997  
C:Accession: S40996  
R:Ainscough, R.  
submitted to the EMBL Data Library, October 1993  
A:Reference number: S40991  
A:Accession: S40996  
A:Molecule type: DNA  
A:Residues: 1-516 <AIN>  
A:Cross-references: EMBL:Z27078; NID:9414627; PID:9414633  
C:Genetics:  
A:Introns: 56/2; 126/1; 152/1; 181/2; 221/2; 255/3; 283/3; 360/1; 396/3; 426/1; 466/3

Query Match 21.4%; Score 413; DB 2; Length 516;  
Best Local Similarity 29.5%; Pred. No. 2.2e-27;  
Matches 103; Conservative 61; Mismatches 137; Indels 48; Gaps 10;  
QY 17 LFGVLG-AATISCRNEEGK-----AVDW-----FTFYKLPKR 47  
Db 171 MFDIIGKSGTICQNIETTHKMLNCLIALNLSLKKFQVSDMQHFFLLFVYKLPKL 230

QY 48 QNKESG---ETCLEVLYLDSTTRSWRKSEQLMNDTKSVLGRITLQOLYEAAYAKSNNTAYL 104  
Db 231 WHPDNPVPSISGTGFLYFVNKNKWLMPQGDVNNNAVYTLQOYIN---SNMNTFSY 287  
QY 105 IYNDGVPKPNVGRKYGHTKGLLMNRVQGFLLHISIPQFPPIPEGYDYPPTGRNGOS 164  
Db 288 WYNDWPDSTIWSNNGSHAKGVTFDQYTGFWLHISIPKFP--SKDMFRFPNAHYGQM 345  
QY 165 GICITFKYNQYEAIDSQLVCNPNVYSCSIPATFHQELIHPQLCTRASSEIPQRLTT 224  
Db 346 GICISYNTVSLATIAQQLFYNTFTYQFNLPQSPANQFPVLSQLKNKEYNKSPPILTSTKV 405  
QY 225 LQSAQOGKFLHFAKSDSFLDDIEAAMQRLKTHLTETWQRK---RELPSNCSLPYHV 281  
Db 406 LKSLGGQHFHFHFAKTGEWGDLYSDVFGFTLKSSIKVETWNHQSGDEYNLPSVCD-PNHV 464  
QY 282 YNKAIKLSRHSY---FSSYQDHAKWCISQKGTNR-----WTCIGDLNR 323  
Db 465 QSTWSAKYIRLPYALDYSYEDHSAKFVYAYSESSSKPPIPVYVCIGDINR 513

RESULT 7  
S44793  
F09G8.2 protein - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 14-Sep-1994 #sequence\_revision 12-May-1995 #text\_change 23-Mar-2001  
C:Accession: S44793  
R:Anderson, K.  
submitted to the EMBL Data Library, February 1993  
A:Description: Sequence of the C. elegans cosmid F09G8.  
A:Reference number: S44796  
A:Accession: S44793  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-238 <AND>  
A:Cross-references: EMBL:L11247; NID:9156280; PID:9156283  
C:Genetics:  
A:Introns: 32/2; 171/3

Query Match 16.5%; Score 319; DB 2; Length 238;  
Best Local Similarity 31.3%; Pred. No. 9.1e-20;  
Matches 76; Conservative 41; Mismatches 106; Indels 20; Gaps 10;

QY 6 LRTSFALLFLGVLGAATISCRNEEGKAVDWTFYKLPKRQKESGETGLELYLDST 65  
Db 1 MRLYFVLIFSVIF-TTGNKGIQCKNMGKSYDWFVVKLPKLSG--AGTSGKEFYFDPE 57  
QY 66 TRSWRKSEQLMNDTKSVLGRITLQOLYEAAYAKSNNTAYLIYNDGVPKPNVSRKY-GHTK 124  
Db 58 SSDWTRGNDI-NDPNVAVGATVSQVYS--ADKSNN-FWFMYS--DDPIKASDSYRGHAK 111  
QY 125 GLILMNRVQGFLLHISIPQFPPIPEGYDYPPTGRNGSGGICITKYNQYEAIDSQLV 184  
Db 112 GVSLFDDTTGFWLHISVNPFPPI--KFSFDPNTAKYQGSFFCASMEVQHUTELAEHWKY 169  
QY 185 CNPNVYSCSIPATFHQELIHPQLCTRASSEIPGR-----LFTLSAQOGKFLHFAK 239  
Db 170 IQATPYIINIPEKY---ATRPFTLKNVEAKQSLPSRATQFWISKPIKTQGVTLMAVAKH 226  
QY 240 DSF 242  
Db 227 KKF 229

RESULT 8  
A44164  
secreted glycoprotein, 43 K - trichina  
C:Species: Trichinella spiralis (trichina)  
C>Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 09-Sep-1997  
C:Accession: A44164; A60630; S27862  
R:Vassilatis, D.K.; Despommier, D.; Misek, D.E.; Polvere, R.I.; Gold, A.M.; Van der Ploeg  
J. Biol. Chem. 267, 18459-18465, 1992

A;Title: Analysis of a 43-kDa glycoprotein from the intracellular parasitic nematode *Trichostrongylus axei*  
A;Reference number: A44164; MUID:92406752; PMID:1382055  
A;Accession: A44164  
A;Molecule type: mRNA  
A;Residues: 1-344 <VAS>  
A;Cross-references: EMBL:M95499; NID:G12534; PTD:G162535  
A;Note: sequence extracted from NCBI backbone (NCBIN:113310, NCBIPI:113321)  
A;Note: part of this sequence was confirmed by protein sequencing  
R;Gold, A.M.; Despommier, D.D.; Buck, S.W.  
Mol. Biochem. Parasitol. 41, 187-196, 1990  
A;Title: Partial characterization of two antigens secreted by L1 larvae of *Trichinella spiralis*  
A;Reference number: A60630; MUID:90377287; PMID:2398916  
A;Accession: A60630  
A;Molecule type: protein  
A;Residues: 23,'X',25-62,'YGSF',66-80,'T',115-116,'X',118-132,249-278 <GOL>  
C;Keywords: glycoprotein

Query Match 6.1%; Score 117; DB 2; Length 344;  
Best Local Similarity 20.7%; Pred. No. 0.029;  
Matches 74; Conservative 61; Mismatches 142; Indels 80; Gaps 20;

QY	13	LFLGLFGV-----LGAATISCRNEEGKAVDWTFYKLPKRNKESGETG-LEVLVYLDLS	64
DB	5	IFLSAFVWIIHNCLOIHAANCTCRATDD-TEWFLLEK-----PVGLLAKAIISP	53

65 TTRSWRKSEQLMN-DTKSVLGRTLQOLYEAYASKSNNTAVLIYNDGVPKVPNYSKYGHT 123

QY

66

67

68

69

70

71

72

73

74

75

76

77

78

79

80

81

82

83

84

85

86

87

88

89

90

91

92

93

94

95

96

97

98

99

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

<

QY  
124 KGLLWNR--VQGFWLIHSIQ-FPPIPEGYDPTGRRNGSQSGICITFKKNOYEADS 180  
110 KYILMECNMGGEMLYRBBADP-NSVASCKETSGHMAICISFSDNVDI-- 165

```
QY      181 QLLVCNPNVSCSIPTAFHOE-LIHPOLCTRAS-----SSEIPGRLL----- 221
```

QY 222 LTTTLSAQOGKFLHFAKSDSFLLDDIFAAWMAORLTHLLTETWORKQELPSNCSLPYHV 281

QY 282 YNKAIKLSRHSYFSYQDHAKWCISQKTKNRWTCIGDLNRSPhQAFFSGFICTQ 338  
dp 267 GFVKVVK-SPVTIDGTGDRSK-----DKSQAWIDD---KPVFCTFTNGSYSTKQ 312

RESULT 9  
 H71525  
 probable outer membrane leader peptide (omp) CT351 - Chlamydia trachomatis (serotype D, H71525)  
 C:Species: Chlamydia trachomatis  
 C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 04-Feb-2000  
 C:Accession: H71525  
 R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, R.  
 Science 282, 754-759, 1998  
 A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis  
 A:Reference number: A71570; MUID:93000809; PMID:9784136  
 A:Accession: H71525  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-697 <ARN>  
 A:Cross-references: GB:AE001308; GB:AE001273; NID:G3328766; PIDN:AAC67946.1; PID:G3328777  
 A:Experimental source: serotype D, strain UW-3/Cx  
 C:Genetics:  
 A:Gene: CT351

Query Match 5.0%; Score 97.5; DB 2; Length 697;  
Best Local Similarity 24.9%; Pred. No. 3.4;  
Matches 62; Conservative 25; Mismatches 109; Indels 53; Gaps 12;

Qy	114	VNYSRK-----YCHTKGLLLWN-RVQGF-WLIH-----SIPPPPIEEGYDPPTG	158
Dy	85	VNYSRKILICDLYEVEDTDSCLLTNGRCSLYPWFICGGTITISRSSIIHKGYLSTSEG	144

RESULT 11

S63059  
hypothetical protein YNL118c - yeast (*Saccharomyces cerevisiae*)

Qy	159	RRNGSGICITFKFNQYEADISQLLVNCPNVYSCISIPATEHQEL-----THMPOLCTRA	211
Db	145	---PQRHICISGDLKYSS-DSVLSNGFSRLSICNTPVLLLPQISIMPEIPKPPITFRG	200
Qy	213	SSSEIPRLITTLQSAQGQKFLHFAKGDSDFLDDIFAAMWAQRLKTHLLTETWQRKQELP	272
Db	201	GSGGFLGSYLVGSYSPISKK---HCSTTILFDLGGFKHGIG-----LGVNRFSSQENP	250
Qy	273	SNCSLPYHYNIKAIKLSRHSYFSSYODHAKWCISQKGTKNRWTCIGDINRSPHOAFRSG	332
Db	251	SN-----AINIK-----SYVAH-RLAIDSSGAKORYRLHGDPTFSKERAHLAG	292
Qy	333	GFICTQNWQ	341
Db	293	EFHLSDSWE	301

RESULT 1.0

C81542  
 conserved hypothetical protein CP0756 [imported] - Chlamydoiphila pneumoniae (strain AR39)  
 C:Species: Chlamydoiphila pneumoniae, Chlamydia pneumoniae  
 C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 11-May-2000  
 C:Accession: C81542  
 R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelber, J.F.; White, O.; Hickey,  
 F.; Reed, T.D.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
 C.; C. Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
 Nucleic Acids Res. 28, 1397-1406, 2000  
 A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.  
 A:Reference number: AB1500; MUID:20150255; PMID:10684935  
 A:Accession: C81542  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-704 <REA>  
 A:Cross-references: GB:AE002234; GB:AE002161; NID:g7189667; PIDN:AAF38559.1; PID:g718967  
 A:Experimental source: strain AR39, HL cells  
 C:Genetics:  
 G:Gene: CP0756

Query Match 5.0%; Score 96.5; DB 2; Length 704;  
Best Local Similarity 19.7%; Pred. No. 4.2;  
Matches 76: Conservative 59; Mismatches 157; Indels 93; Gaps 19;

QY 4 RLHETSFALLFLGFLGVLGNAATISCRNEEG--KAVDNWFTFYKLPKRONKESGSETGLE--Y 59

ph 8 RIMKRCF--LFLASFLVIGSSADALTHOEAVKKQKSYLSHFK-----SVSGIVTIEDGV 59

[illegible]

QY 115 NYSRKYGHTKGLLWN---RVQGFWLHSIPQFPP---IPREGYDYPPTGRNRQGSQGI 168  
159 YNRDQCTIYNCRSMAYWYRCCSMITWTETETIVTQGYTSMSEGGK---KDI.CI. 159  
160 YNRDQCTIYNCRSMAYWYRCCSMITWTETETIVTQGYTSMSEGGK---KDI.CI. 159

QY 169 TFKYQYEAIQSLLVCNPNVYSCSIPATHQEL-----IHPQLCTRASSEIPGRLL 2222

```
QY      223  TTLSAQGQKFLHPAKSDSFLDDIIFAAWMAQLKTHLLTETWQRKQLPNSCLPHVHY 282
QY      223  TTLSAQGQKFLHPAKSDSFLDDIIFAAWMAQLKTHLLTETWQRKQLPNSCLPHVHY 282
```

```

20
210
220
230
240
250
260
270
280
290
300
310
320
330
340
350
360
370
380
390
400
410
420
430
440
450
460
470
480
490
500
510
520
530
540
550
560
570
580
590
600
610
620
630
640
650
660
670
680
690
700
710
720
730
740
750
760
770
780
790
800
810
820
830
840
850
860
870
880
890
900
910
920
930
940
950
960
970
980
990
1000
1010
1020
1030
1040
1050
1060
1070
1080
1090
1100
1110
1120
1130
1140
1150
1160
1170
1180
1190
1200
1210
1220
1230
1240
1250
1260
1270
1280
1290
1300
1310
1320
1330
1340
1350
1360
1370
1380
1390
1400
1410
1420
1430
1440
1450
1460
1470
1480
1490
1500
1510
1520
1530
1540
1550
1560
1570
1580
1590
1600
1610
1620
1630
1640
1650
1660
1670
1680
1690
1700
1710
1720
1730
1740
1750
1760
1770
1780
1790
1800
1810
1820
1830
1840
1850
1860
1870
1880
1890
1900
1910
1920
1930
1940
1950
1960
1970
1980
1990
2000
2010
2020
2030
2040
2050
2060
2070
2080
2090
2100
2110
2120
2130
2140
2150
2160
2170
2180
2190
2200
2210
2220
2230
2240
2250
2260
2270
2280
2290
2300
2310
2320
2330
2340
2350
2360
2370
2380
2390
2400
2410
2420
2430
2440
2450
2460
2470
2480
2490
2500
2510
2520
2530
2540
2550
2560
2570
2580
2590
2600
2610
2620
2630
2640
2650
2660
2670
2680
2690
2700
2710
2720
2730
2740
2750
2760
2770
2780
2790
2800
2810
2820
2830
2840
2850
2860
2870
2880
2890
2900
2910
2920
2930
2940
2950
2960
2970
2980
2990
3000
3010
3020
3030
3040
3050
3060
3070
3080
3090
3100
3110
3120
3130
3140
3150
3160
3170
3180
3190
3200
3210
3220
3230
3240
3250
3260
3270
3280
3290
3300
3310
3320
3330
3340
3350
3360
3370
3380
3390
3400
3410
3420
3430
3440
3450
3460
3470
3480
3490
3500
3510
3520
3530
3540
3550
3560
3570
3580
3590
3600
3610
3620
3630
3640
3650
3660
3670
3680
3690
3700
3710
3720
3730
3740
3750
3760
3770
3780
3790
3800
3810
3820
3830
3840
3850
3860
3870
3880
3890
3900
3910
3920
3930
3940
3950
3960
3970
3980
3990
4000
4010
4020
4030
4040
4050
4060
4070
4080
4090
4100
4110
4120
4130
4140
4150
4160
4170
4180
4190
4200
4210
4220
4230
4240
4250
4260
4270
4280
4290
4300
4310
4320
4330
4340
4350
4360
4370
4380
4390
4400
4410
4420
4430
4440
4450
4460
4470
4480
4490
4500
4510
4520
4530
4540
4550
4560
4570
4580
4590
4600
4610
4620
4630
4640
4650
4660
4670
4680
4690
4700
4710
4720
4730
4740
4750
4760
4770
4780
4790
4800
4810
4820
4830
4840
4850
4860
4870
4880
4890
4900
4910
4920
4930
4940
4950
4960
4970
4980
4990
5000
5010
5020
5030
5040
5050
5060
5070
5080
5090
5100
5110
5120
5130
5140
5150
5160
5170
5180
5190
5200
5210
5220
5230
5240
5250
5260
5270
5280
5290
5300
5310
5320
5330
5340
5350
5360
5370
5380
5390
5400
5410
5420
5430
5440
5450
5460
5470
5480
5490
5500
5510
5520
5530
5540
5550
5560
5570
5580
5590
5600
5610
5620
5630
5640
5650
5660
5670
5680
5690
5700
5710
5720
5730
5740
5750
5760
5770
5780
5790
5800
5810
5820
5830
5840
5850
5860
5870
5880
5890
5900
5910
5920
5930
5940
5950
5960
5970
5980
5990
6000
6010
6020
6030
6040
6050
6060
6070
6080
6090
6100
6110
6120
6130
6140
6150
6160
6170
6180
6190
6200
6210
6220
6230
6240
6250
6260
6270
6280
6290
6300
6310
6320
6330
6340
6350
6360
6370
6380
6390
6400
6410
6420
6430
6440
6450
6460
6470
6480
6490
6500
6510
6520
6530
6540
6550
6560
6570
6580
6590
6600
6610
6620
6630
6640
6650
6660
6670
6680
6690
6700
6710
6720
6730
6740
6750
6760
6770
6780
6790
6800
6810
6820
6830
6840
6850
6860
6870
6880
6890
6900
6910
6920
6930
6940
6950
6960
6970
6980
6990
7000
7010
7020
7030
7040
7050
7060
7070
7080
7090
7100
7110
7120
7130
7140
7150
7160
7170
7180
7190
7200
7210
7220
7230
7240
7250
7260
7270
7280
7290
7300
7310
7320
7330
7340
7350
7360
7370
7380
7390
7400
7410
7420
7430
7440
7450
7460
7470
7480
7490
7500
7510
7520
7530
7540
7550
7560
7570
7580
7590
7600
7610
7620
7630
7640
7650
7660
7670
7680
7690
7700
7710
7720
7730
7740
7750
7760
7770
7780
7790
7800
7810
7820
7830
7840
7850
7860
7870
7880
7890
7900
7910
7920
7930
7940
7950
7960
7970
7980
7990
8000
8010
8020
8030
8040
8050
8060
8070
8080
8090
8100
8110
8120
8130
8140
8150
8160
8170
8180
8190
8200
8210
8220
8230
8240
8250
8260
8270
8280
8290
8300
8310
8320
8330
8340
8350
8360
8370
8380
8390
8400
8410
8420
8430
8440
8450
8460
8470
8480
8490
8500
8510
8520
8530
8540
```

QY 322 NRSPHQAFRSGGFICTONWQIYQAF 346

DD 323 NIGIRV-----DCIMKDNFFES1 341

RESULT 11

RESULT 11

S63059  
hypothetical protein YNL118c - yeast (*Saccharomyces cerevisiae*)

N;/Alternate names: hypothetical protein N1917  
C;/Species: Saccharomyces cerevisiae  
C;/Date: 27-Apr-1996 #sequence\_revision 03-May-1996 #text\_change 26-May-2000  
C;/Accession: S63059; S59701; S67340  
R;/De Antoni, A.; D Angelo, M.; Dal Pero, F.; Sartorello, F.; Pandolfo, D.; Lanfranchi, G.  
submitted to the Protein Sequence Database, April 1996  
A;/Reference number: S63047  
A;/Accession: S63059  
A;/Molecule type: DNA  
A;/Residues: 1-970 <DEA>  
A;/Cross-references: EMBL:Z71394; NID:g1302044; PID:e239800; PID:g1302045; MIPS:YNL118C  
A;/Experimental source: strain S288C  
R;/Tzagoloff, A.A.  
submitted to the EMBL Data Library, June 1995  
A;/Description: Suppressor of a yeast pet mutant.  
A;/Reference number: S59701  
A;/Accession: S59701  
A;/Molecule type: DNA  
A;/Residues: 1-424, 'L', 426-970 <TZA>  
A;/Cross-references: EMBL:L43065; NID:g870733; PID:g870734  
A;/Experimental source: strain D273-10B  
R;/D'Antoni, A.; d'Angelo, M.; dal Pero, F.; Sartorello, F.; Pandolfo, D.; Lanfranchi, G.  
submitted to the EMBL Data Library, February 1996  
A;/Description: The DNA sequence of cosmid 14-13b from chromosome XIV of *Saccharomyces cerevisiae*.  
A;/Reference number: S67327  
A;/Accession: S67340  
A;/Molecule type: DNA  
A;/Residues: 1-970 <DAN>  
A;/Cross-references: EMBL:Z69382; NID:g1183941; PID:e221828; PID:g1183955  
C;/Genetics:  
A;/Gene: SGD:PSU1  
A;/Cross-references: SGD:S0005062; MIPS:YNL118C  
A;/Map position: 14L  
C;/Superfamily: unassigned mutt domain proteins; mutt domain homology  
C;/129-163/Domain: mutt domain homology <MUTT>

Query Match	4.9%;	Score 95;	DB 2;	Length 970;		
Best Local Similarity	22.5%;	Pred. No. 8.6;				
Matches	64;	Conservative 35;	Mismatches 99;	Indels 86;	Gaps 15	
QY	17	LFGVIGNAAT	-----SCRNEEGKAVDWF	TFYKLPKRONKSGG	TGLEYLYLDSTTR---S 68	
DB		: : : : :	: : : : :	: : : : :	: : : : :	
QY	179	IFLISGVSEVF	NFKPQVRNEIDK-I	IEWDFKIKSTWYK---	SNIKYLLINSMRPLSM 233	
DB		: : : : :	: : : : :	: : : : :	: : : : :	
QY	69	W-RKSQLMND	-----TKSVLGR	T-----LQOIYEAV	AKSNNTAYLIY 106	
DB		: : : : :	: : : : :	: : : : :	: : : : :	
QY	234	WLHQIQIKNED	OLKSYABEQKLLG	ITKEEQIDPCREL	LNMLHTAVQANSNNA--VS 291	
DB		: : : : :	: : : : :	: : : : :	: : : : :	
QY	107	NDGVKXP	-----	VNYSRKYGHTKGL	-----127	
DB		: : : : :	: : : : :	: : : : :	: : : : :	
QY	292	NGQVPSSQEL	QHLKQSGEHNQK	QDQSSFSQQQPS	IFPSSLEPFFANNKNV	IPPTWEMA 351
DB		: : : : :	: : : : :	: : : : :	: : : : :	
QY	128	--LWNVQGF	WLHISIPQFPPI	PEEGYDYPTGR	NRNGQSGICITFKYN	QVEAIDSQLLV 185
DB		: : : : :	: : : : :	: : : : :	: : : : :	
QY	352	NVFMSPNPQ	LATWNGQP-FAPFP	---FMLPLTNNS	ANSANPIPTVPFN-F	NAENPNMAFG 406
DB		: : : : :	: : : : :	: : : : :	: : : : :	
QY	186	NNVYSCSIP	ATHQELIHPQL	CTRAS---SSEIP	GRLLTLOS 227	
DB		: : : : :	: : : : :	: : : : :	: : : : :	
QY	407	VFNHNLSPG	AV-SQPSFLP	APLPDQSGYSSSP	GGLOLDLNS 449	
DB		: : : : :	: : : : :	: : : : :	: : : : :	

RESULT 12  
B71621  
probable membrane associated protein PFB0190c - malaria parasite (Plasmodium falciparum)  
C:Species: Plasmodium falciparum  
C:Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 21-Jul-2000  
C:Accession: B71621  
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;  
Peters, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.  
Science 282, 1126-1132, 1998  
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.  
A:Reference number: A71600; MID:99021743; PMID:9804551  
A:Accession: B71621

A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-2295 <GAR>  
A:Cross-references: GB:AE001379; GB:AE001362; NID:g3845118; PIDN:AACT71827.1; PID:g384511  
A:Experimental source: clone 3D7  
C:Genetics:  
A:Gene: PFB0190C

Query Match 4.9%; Score 94.5; DB 2; Length 2295;  
Best Local Similarity 24.9%; Pred. No. 29;  
Matches 44; Conservative 25; Mismatches 79; Indels 29; Gaps 8;

Qy 7 RTSFALLFLGVLGAATISCRNBEGKAVDWFTFYKLPKRNKESGETGLELYLSTT 66  
Db 1089 RNSYALAALGIYIYIGNESIGIERDEIKA--FEFWKKAADQGDTTTSAISTGYAYLDEYK 1145  
Qy 67 RSWKSEQLMN-DTKSVLGRITQQLYEAYASNSNNTAYLIYNDGVKPKVNSRYK----- 120  
Db 1146 KFLKEELVKNMDREDIIL--TWIHLNSTYKDKKNVTLEMPQESSEKKNQKKKKEKKEQD 1203  
Qy 121 GHTKGLLLNWRVQGFWLHISI-----PQPPPIPE-----EGYDYPPTGRRNGQS 164  
Db 1204 GNTDG----DRVDD-KIVQNVGNVQSYGNVDSEMGRRNGSIDGFSNPPSGGLNNVS 1255

RESULT 13  
DEFFG6  
glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) - fruit fly (*Drosophila melanogaster*)  
N/Alternate names: D-glucose-6-phosphate:NADP+ oxidoreductase  
C/Species: *Drosophila melanogaster*  
C/Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 11-Jun-1999  
C/Accession: JT0272; A38006  
R/Fouts, D.; Ganguly, R.; Gutierrez, A.G.; Lucchesi, J.C.; Manning, J.E.  
Gene 63, 261-275, 1988  
A/Title: Nucleotide sequence of the *Drosophila* glucose-6-phosphate dehydrogenase gene and  
A/Reference number: JT0272; MUID:88255872; PMID:2838391  
A/Accession: JT0272  
A:Molecule type: genomic RNA  
A:Residues: 1-523 <FOU>  
A:Cross-references: GB:M26674; NID:g157468; PIDN:AAA51463.1; PID:g157470  
A:Accession: A38006  
A:Molecule type: mRNA  
A:Residues: 1-523 <FOU2>  
A:Cross-references: GB:M26674; NID:g157468; PIDN:AAA51463.1; PID:g157470  
A/Note: the authors translated the codon AAG for residue 214 as Asn  
C/Comment: This is the first and dominant regulatory enzyme in the hexose monophosphate  
C/Genetics:  
A:Gene: FlyBase:zw  
A:Cross-references: FlyBase:FBgn0004057  
A:Introns: 6/2; 94/2; 166/2  
C/Superfamily: Glucose-6-phosphate dehydrogenase  
C/Keywords: oxidoreductase; pentose phosphate pathway  
F/209/Active site: Lys #status predicted

Query Match	4.9%;	Score 94;	DB 1;	Length 523;
Best Local Similarity	18.3%;	Pred. No. 4.7;		
Matches	50;	Conservative 43;	Mismatches 86;	Indels 94;
				Gaps 12;
QY	101	TAVLIYNDG-VKPVNYSKYGHTKGLLWNRVQGFWLIIHSIOFFPIPEGVDYPP---	156	
Db	55	TLWWLVRDDLPLPTKFC---GYRSMLTVDLSIKE---OCLPYMKVQPFHQEKYEEFWA	107	
QY	157	-----TGRNGSGSICITFKYQNYEADISQLLY-----CNPNVYSCSIPATFHQELIHM	205	
Db	108	LNEYVSGRYDGRGTG-----FELLNQOLETMENKKNRIFYLALPFSVPEEVTVNI	158	
QY	206	POLCTR-----ASSSEIEGRLTLTLOSAGQGFILHFAKSDSFLDDIFA	248	
Db	159	KQICMSVCGWNRVIEKPGRRDASAGGSDHLAGLFQEDQLFRIDHYLQKE-----	210	
QY	249	AWMAQSLKT-----HLITETWOKROELPNSCLPYHVYNKAIKLS-----RHSY	294	
Db	211	---MVQKLMITIRFGNKILSTWNRE-----NIAVLITFKPEPFGTQGRGGY	253	

QY 295 FSSY-----QDHAKWCISQKGTKNRTICGD 320  
 Db 254 FDEFGIIRDVMQNHLLQILSLVAMEKPVSCHPD 286

RESULT 14  
 E81286  
 probable polysaccharide modification protein Cj1413c [imported] - Campylobacter jejuni  
 C:Species: Campylobacter jejuni  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 03-Jun-2002  
 C:Accession: E81286  
 R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling  
 C.W.; Quail, J.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrell  
 Nature 403, 665-668, 2000  
 A>Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp  
 A:Reference number: A81250; MUID:20150912; PMID:10688204  
 A:Accession: E81286  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-394 <PAR>  
 A:Cross-references: GB:AL139078; GB:AL111168; NID:G6968723; PIDN:CAB73837.1; PID:G696884  
 A:Experimental source: serotype O2, strain NCTC 11168  
 C:Genetics:  
 A:Gene: Cj1413c  
 C:Superfamily: kpsD protein

Query Match 4.8%; Score 93.5; DB 2; Length 394;  
 Best Local Similarity 20.3%; Pred. No. 3.6;  
 Matches 74; Conservative 39; Mismatches 100; Indels 151; Gaps 17;

QY 43 KLPKQNK--ESGETGLEVLYLDSTTRSWKSEQLMNDTKSVLGRTLQOLYEAYASKNN 100  
 Db 33 KVRKQTKVKLFNFNGDFFFPVSGTRC--KQDE-----KQLENFYRDFQNKKI 80

QY 101 TAYLIYNDGVKPVNYSRYGHTK-----GLLLNWRVQGFWLIHSIPQPPPIPEEGY 152  
 Db 81 DAILMYND-----CRILIAKAIKVAKELGIEIW-----IFEEGY 114

QY 153 DYPPTGRRNQSGICITFKYQVYEAIDSQLVCNPNVYSCSPATPHQLIHPQLCTRA 212  
 Db 115 LRP-----YCTLEKDGVNANS--LPRDKNFYLSQ-----NIFTKE 149

QY 213 SSSETPGRLTLTQSAQOKFLHPAKSDSLDDIFA-----AW 250  
 Db 150 SIKEIPG-----GKFWAF--DAFLVWLFAPILAPPFNKLIHRTLYPPFFLFW 196

QY 251 MAQRKTHLLTETWQRKQ-----ELPSNCSLPYHYVYNKAIKLSRHSYFS 296  
 Db 197 FRSLYEKYLYKTEKLNKINLEKKYFALILQVYSDTQIKYH-YK-KSIEHFIEITIL 254

QY 297 SYQDHA-----KWCISQKGTKNRTICGIDLNRSFHOAFRSGGFICTQNWQIYQAFQ 348  
 Db 255 SPANHARAKSYLVFKHHPMDRGYKNYFKLINDLSRKYH-----VEG 295

QY 349 LVLY 352  
 Db 296 RVLY 299

RESULT 15  
 F82873  
 hypothetical protein UUS78 [imported] - Ureaplasma urealyticum  
 C:Species: Ureaplasma urealyticum  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 15-Sep-2000  
 C:Accession: F82873  
 R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
 submitted to GenBank, February 2000  
 A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min  
 A:Reference number: A82870  
 A:Accession: F82873  
 A>Status: preliminary  
 A:Molecule type: DNA

A:Residues: 1-526 <GLA>  
 A:Cross-references: GB:AE002156; GB:AF222894; NID:G6899580; PIDN:AAF3092.1; GSPDB:GN001  
 A:Experimental source: serovar 3; biovar 1  
 C:Genetics:  
 A:Gene: UUS78  
 A:Genetic code: SGC3  
 C:Superfamily: Ureaplasma urealyticum hypothetical protein UUS78

Query Match 4.8%; Score 93; DB 2; Length 526;  
 Best Local Similarity 19.4%; Pred. No. 5.7; Indels 136; Gaps 18;  
 Matches 84; Conservative 58; Mismatches 155;

QY 4 RLRTSFALLFTGLFGVLGAATISCRN-----EERG--AVDM 38  
 Db 6 KIITSSALLFTGVATISIAACSKQNNQKTKNFNSQEQANKIIEAILNKKEGKONFIDW 65

QY 39 FFYKLPKRONKESGETGLEVLYLDSTTRSWKSEQLMNDTKSVLG-----RTLOQLY 91  
 Db 66 VHMNSPFPQARKLAK-DLD-KYIDLTKSKNFKSMIDKNEYKSTFSAQIENVIRIHKDI 123

QY 92 EAYAS-----KSNNTAYLIYNDGVKPVNYSRYGHTKGLLLNWRVQGFWLIHSIPQ 143  
 Db 124 KMYANSPFWKWRQKKTALFLINYA---PIQKDSQINAPGILL---PAEYPLLYSKED 177

QY 144 FPPPIPEEGYDYPPTGRRNG-----QSGICITFKYQY-----AIDSO 181  
 Db 178 FEELPGFGAFTPTPKSMDAVALDLDTWKSGSLSSKGSVEKSVLASKLQNSFAKTADKV 237

QY 182 LLVCN-----PNVY-----SCSIPATF-----HQLIHPQLC 209  
 Db 238 VVIYDDAIVPNTYDKNGNRNLKITOFFANWMINQNYKYLAREBFLKINOKNDLIPLMST 297

QY 210 TRASSEIPG--RLTLTQSAQOKFLHPAKSDSLDDIFAAMWAQRKTHLLTETWQRK 267  
 Db 298 VHASVGIIGMERNLYTLKAFGMP-----KDELEHLKAQEKFKIPTLKLLTND 347

QY 268 ROELPSNCSLPYHYVYNKAIKLSRHSYFSYQDHAK--WCISQKGTKNRTIC-- 318  
 Db 348 ELELKN-----GIQTIKTGIDSPFKYHRDQNRDDWKIATNSQVLDICIALGLKPD 398

QY 319 ----GDLNRSFHO 327  
 Db 399 LLVKGELSTSVHE 411

Search completed: October 14, 2004, 00:32:23  
 Job time : 31.1224 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 22:14:19 ; Search time 24.6034 Seconds  
(without alignments)  
755.549 Million cell updates/sec

Title: US-10-790-589-4

Perfect score: 1933

Sequence: 1 MWARLLRTSFALLFLGLFGV.....QNWQIYQAFQGLVLYESCK 357

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 segs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	613	31.7	360	1	DRN2_HUMAN	O00115 homo sapien
2	599	31.0	353	1	DRN2_MOUSE	P56542 mus musculu
3	593	30.7	364	1	DRN2_PIG	O62855 sus scrofa
4	505	26.1	378	1	YMW6_CAEEL	P34508 caenorhabdi
5	490.5	25.4	375	1	YKUS_CAEEL	O17778 caenorhabdi
6	472.5	24.4	367	1	YLS2_CAEEL	P34387 caenorhabdi
7	459	23.7	276	1	DRN2_BOVIN	P56541 bos taurus
8	95	4.9	704	1	NEUL_HUMAN	Q9BYT8 homo sapien
9	95	4.9	970	1	PSUL_YEAST	P53550 saccharomyc
10	93.5	4.8	781	1	APR2_SULTO	Q974N6 sulfolobus
11	93	4.8	6669	1	NEBU_HUMAN	P20929 homo sapien
12	92	4.8	524	1	GSPD_DROME	P12646 drosophila
13	92	4.8	1409	1	HAP1_HAETN	P44596 haemophilus
14	91.5	4.7	305	1	Y348_MYCPN	P52555 mycoplasma
15	91	4.7	1025	1	ITR8_HUMAN	P53708 homo sapien
16	90.5	4.7	521	1	PGS1_SACFS	P79001 saccharomyc
17	89	4.6	557	1	CNE6_HUMAN	O95741 homo sapien
18	89	4.6	557	1	CNE6_MOUSE	O92140 mus musculu
19	88.5	4.6	1456	1	MANR_HUMAN	P22897 mus sapien
20	88	4.6	424	1	VK04_VACCV	P18377 vaccinia vi
21	88	4.6	424	1	CUI8_HUMAN	Q9NVD3 homo sapien
22	88	4.6	641	1	TERL_LAMBD	P03708 bacterioph
23	87.5	4.5	358	1	Y4EE_RHISN	P55428 rhizobium s
24	87	4.5	685	1	SC31_HUMAN	Q07837 homo sapien
25	86	4.4	704	1	NEUL_RABIT	P42675 cryptocolag
26	85.5	4.4	508	1	TRPE_SYNY3	P20170 synecocyst
27	85.5	4.4	521	1	FGS1_YEAST	P25578 saccharomyc
28	85.5	4.4	587	1	SKU5_ARATH	Q9SU40 arabidopsi
29	85.5	4.4	775	1	ECEL_RAT	O9JHL3 rattus norv
30	85.5	4.4	951	1	SVV_ECOLI	P07118 escherichia
31	85.5	4.4	954	1	SVV_HAETN	P43834 haemophilus
32	85	4.4	386	1	SAVI_MOUSE	Q8VEB2 mus musculu
33	85	4.4	387	1	GSPL_ECOLI	P45763 escherichia

34	85	4.4	433	1	UROK_MOUSE	P06869 mus musculu
35	85	4.4	655	1	SCAG_MOUSE	Q9WU39 mus musculu
36	85	4.4	704	1	NEUL_PIG	Q02038 sus scrofa
37	85	4.4	833	1	CNSA_RAT	O54735 rattus norv
38	85	4.4	880	1	LYTD_BACSU	P19848 bacillus su
39	84.5	4.4	399	1	GLF_MYCPN	P75499 mycoplasma
40	84.5	4.4	426	1	Y958_METJA	Q58368 methanococc
41	84.5	4.4	557	1	ALO_CANAL	O93852 candida alb
42	84	4.3	383	1	SAVI_HUMAN	Q9H4B6 homo sapien
43	84	4.3	475	1	RP54_THIFE	P24695 thriobacillu
44	84	4.3	1458	1	PA2R_RABIT	P49260 oryctolagus
45	83.5	4.3	1256	1	SA3_RAT	Q99M76 rattus norv

## ALIGNMENTS

RESULT 1

DRN2_HUMAN	AC	O00115	043910;	STANDARD;	PRT;	360 AA.
DT	15-JUL-1998	(Rel. 36, Created)				
DT	15-JUL-1998	(Rel. 36, Last sequence update)				
DT	10-OCT-2003	(Rel. 42, Last annotation update)				
DE	Deoxyribonuclease II precursor (EC 3.1.22.1) (DNase II) (Acid DNase)					
DE	(Lysosomal DNase II) (R31240_2).					
GN	DNASE2 OR DNL2.					
OS	Homo sapiens (Human).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
OX	NCBI_TaxID=9606;					
RN	[1]	SEQUENCE FROM N.A.				
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=98112802; PubMed=9446563;					
RA	Yasuda T., Takeshita H., Iida R., Nakajima T., Hosomi O.,					
RA	Nakashima Y., Kishi K.;					
RT	"Molecular cloning of the cDNA encoding human deoxyribonuclease II.";					
RL	J. Biol. Chem. 273:2610-2616(1998).					
RN	[2]	SEQUENCE FROM N.A.				
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=20363736; PubMed=10903447;					
RA	Krieser R.J., Eastman A.;					
RT	"Deoxyribonuclease II: structure and chromosomal localization of the					
RT	murine gene, and comparison with the genomic structure of the human					
RT	and three C. elegans homologs.";					
RL	Gene 252:155-162(2000).					
RN	[3]	SEQUENCE FROM N.A.				
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=98382521; PubMed=9714827;					
RA	Baker K.P., Baron W.F., Henzel W.J., Spencer S.A.;					
RT	"Molecular cloning and characterization of human and murine DNase					
RT	II.";					
RL	Gene 215:281-289(1998).					
RN	[4]	SEQUENCE FROM N.A.				
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=98321218; PubMed=9647784;					
RA	Shiohara D., Tanuma S.-I.;					
RT	"Cloning of cDNAs encoding porcine and human DNase II.";					
RT	Biochem. Biophys. Res. Commun. 247:864-869(1998).					
RN	[5]	SEQUENCE FROM N.A.				
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=99123683; PubMed=9924608;					
RA	Yasuda T., Takeshita H., Iida R., Tsutsumi S., Nakajima T., Hosomi O.,					
RA	Nakashima Y., Mori S., Kishi K.;					
RT	"Structure and organization of the human deoxyribonuclease II (DNase					
RT	II) gene.";					
RN	Ann. Hum. Genet. 62:299-305(1998).					
RP	SEQUENCE FROM N.A.					
RA	Lamerdin J.B., McCready P.M., Stilwagen S., Ramirez M., Carrano A.;					
RT	"Characterization by genomic sequence analysis of a gene-rich 111 kb					
RT	region of 19p13.2 containing the human DNA repair gene, RAD23A.";					
RT	Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.					

[7]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Muscle;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
[8]  
RP CARBOHYDRATE-LINKAGE SITE ASN-212.  
RX MEDLINE=22660472; PubMed=12754519;  
RA Zhang H., Li X.-J., Martin D.B., Aebersold R.;  
RT "Identification and quantification of N-linked glycoproteins using  
hydrazone chemistry, stable isotope labeling and mass spectrometry";  
RL Nat. Biotechnol. 21:660-666 (2003).  
CC -!- FUNCTION: Hydrolyzes DNA under acidic conditions with a preference  
for double-stranded DNA. Has a possible role in apoptosis.  
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to nucleoside 3'-  
phosphates and 3'-phosphooligonucleotide end-products.  
CC -!- SUBCELLULAR LOCATION: Lysosomal.  
CC -!- PTM: Glycosylated.  
CC -!- SIMILARITY: Belongs to the DNase II family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AB004574; BAR28623.1; -  
CC EMBL; AF047016; AAC77366.1; -  
CC EMBL; AF045937; AAC35751.1; -  
CC EMBL; AF060222; AAC39852.1; -  
CC EMBL; AB008564; BAB55598.1; -  
CC EMBL; AB008559; BAB55598.1; JOINED.  
CC EMBL; AB008560; BAB55598.1; JOINED.  
CC EMBL; AB008561; BAB55598.1; JOINED.  
CC EMBL; AB008562; BAB55598.1; JOINED.  
CC EMBL; AB008563; BAB55598.1; JOINED.  
CC EMBL; AD000092; AAB51172.1; ALT\_SEQ.  
CC EMBL; BC010419; AAH10419.1; -  
CC PIR; J02026; J02026.  
CC Genew; HGNC:2960; DNASE2.  
CC MIM; 126350; -  
CC GO; GO:0005764; C:lysosome; TAS.  
CC GO; GO:0004531; F:deoxyribonuclease II activity; TAS.  
CC GO; GO:0003677; F:DNA binding; TAS.  
CC GO; GO:0006259; P:DNA metabolism; TAS.  
CC InterPro; IPR004947; DNase\_II.  
CC Pfam; PF02265; DNase\_II; 1.  
KW Hydrolyase, Endonuclease, Nuclease, Lysosome, Zymogen, Signal;  
KW Apoptosis, Glycoprotein, Polymorphism.  
FT SIGNAL 1 18 POTENTIAL.  
FT PROPEP 19 106 BY SIMILARITY.  
FT CHAIN 107 360 DBOXYRIBONUCLEASE II.

FT CARBOHYD 86 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 212 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 266 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 290 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARIANT 314 R->L (in dbSNP:1061192).  
FT /FTID=VAR\_012044.  
SQ SEQUENCE 360 AA; 39581 MW; DFIBBFBA9676EA CRC64;  
Query Match 31.7%; Score 613; DB 1; Length 360;  
Best Local Similarity 36.9%; Pred No. 2,1e-45;  
Matches 129; Conservative 61; Mismatches 148; Indels 12; Gaps 6;  
QY 13 LFLGFLGVLGAATISCRNEEGKAVDWTFYKLPK-RQNKESGETGLEIYLYDSTRWRK 71  
Db 4 LLLAALLCPAGALTCYDGGQVDFWVYKLLPALRGSGEAAQRLQYKYLDESSGMRD 63  
QY 72 SSQLMNDYKSVLGRLOIYEAYASKNNATVLIYNDGVKPK--VNYSRKYHTKGLILW 129  
Db 64 GRALINSPGAVGRSLQPLR---SNTSOLAFLIYNDQPPQSKAQDSSMRGHTKGVLL 120  
QY 130 NRVQGFHLIHSIPQF--PIPERGYDYPPTGRNGSGGICITFYKQYEAID9QLLYCNP 188  
Db 121 DHGGFWLHVHVPNPPPPASSAAYSWPHSACTYGTLLCVSPFFAQFSKMGKOLTYEPW 180  
QY 189 VYCSIPATFHQELIHPOLCTRASSEIPGRLLITLQSAQOKFLHFAKSDSFLDDIYA 248  
Db 181 VINYQLEGIFAQEFPPDLENVWKGHVHVSQEFWNSITLTSQAGAVFQSFKFGDDLYS 240  
QY 249 AWAQRLKTHLTETWQRKQELPNSCLPYHYVNIKAIKLSRHS--YFSSYODHAKWCI 306  
Db 241 GWLAAALGTNLQVQFWHKTGVLPSNCSDIQWLVNQLAFPGPAGPFSFNSTHDSKWCV 300  
QY 307 SOKGTNRNWTCTGDLNRSRPHQAFRSGGFTCTQNWQIYQAFQGLVLYESC 356  
Db 301 SPKGP---WTCVGMNMQEGRGGGTLCALPALWKAQPLVKNYQPC 347  
RESULT 2  
ID DN2 MOUSE STANDARD; PRT; 353 AA.  
AC P56542; O55053;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Deoxyribonuclease II precursor (EC 3.1.22.1) (DNase II) (Acid DNase)  
DE (Lysosomal DNase II).  
DE DNASE2 OR DNASE2A OR DN12.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98382521; PubMed=9714827;  
RA Baker K.P., Baron W.F., Henzel W.J., Spencer S.A.;  
RA "Molecular cloning and characterization of human and murine DNase  
II.";  
RL Gene 215:281-289 (1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20363736; PubMed=10903447;  
RA Krieser R.J., Eastman A.;  
RA "Deoxyribonuclease II: structure and chromosomal localization of the  
murine gene, and comparison with the genomic structure of the human  
and three C. elegans homologs.";  
RL Gene 252:155-162 (2000).  
CC -!- FUNCTION: Hydrolyzes DNA under acidic conditions with a preference  
for double-stranded DNA. Has a possible role in apoptosis.  
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to nucleoside 3'-  
phosphates and 3'-phosphooligonucleotide end-products.  
CC -!- SUBCELLULAR LOCATION: Lysosomal.  
CC -!- SIMILARITY: Belongs to the DNase II family.



AC	P34508; Q9NRAQ0; Q9NCFF8;
AD	01-FEB-1994 (Rel. 28, Created)
AE	28-FEB-2003 (Rel. 41, Last sequence update)
AF	28-FEB-2003 (Rel. 41, Last annotation update)
AG	Hypothetical protein K04H4.6 in chromosome III precursor.
AH	K04H4.6.
AI	Caeenorhabditis elegans.
AJ	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea;
AK	Rhabditiidae; Peloderinae; Caenorhabditis.
AL	NCB1_TaxID=6239;
AM	[1]
AN	SEQUENCE FROM N.A.
AO	STRAIN=Bristol N2;
AP	MEDLINE=94150718; PubMed=7906398;
AQ	Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
AR	Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fraser A.,
AS	Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
AT	Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
AV	Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
AW	Lafrelle P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
AX	Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
AY	Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
AZ	Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
BA	Waterson R., Watson A., Weinstein L., Wilkinson-Sproat J.,
BB	Wohlman P.;
BC	"2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
BD	elegans.";
BE	Nature 368:32-38(1994).
BF	[2]
BG	REVIEWS, AND ALTERNATIVE SPLICING.
BH	Durbin R.;
BI	Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
BJ	[3]
BK	SEQUENCE OF 18-369 FROM N.A. (ISOFORM A).
BL	MEDLINE=20363736; PubMed=10903447;
BM	Krieger R.J., Eastman A.;
BN	"Deoxyribonuclease II: structure and chromosomal localization of the
BO	murine gene, and comparison with the genomic structure of the human
BP	and three C. elegans homologs.";
BQ	Gene 252:155-162(2000).
BR	-!- ALTERNATIVE PRODUCTS:
BS	Event=Alternative splicing; Named isoforms=2;
BT	Name=b;
BU	IsoId=P34508-1; Sequence=Displayed;
BV	Name=a;
BW	IsoId=P34508-2; Sequence=VSP_001304;
BX	Notes=No experimental confirmation available;
BY	-!- SIMILARITY: Belongs to the DNase II family.
BZ	[1]
CA	This SWISS-PROT entry is copyright. It is produced through a collaboration
CB	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CD	use by non-profit institutions as long as its content is in no way
CE	modified and this statement is not removed. Usage by and for commercial
CF	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CG	or send an email to license@isb-sib.ch).
CH	EMBL; Z19153; CAD45582.1; --
CI	DR EMBL; Z27078; CAD45582.1; JOINED.
CJ	DR EMBL; Z27078; CAAB1586.2; --
CK	DR EMBL; Z19153; CAAB1586.2; JOINED.
CL	DR EMBL; Z27078; CAB76943.1; --
CM	DR EMBL; AF220525; RAF43008.1; --
CN	DR FIr; S40996; S40996.
CO	DR WormPep; K04H4.6a; CE25043.
CP	DR WormPep; K04H4.6b; CE31813.
CQ	DR InterPro; IPR004947; DNase_II.
CR	DR Pfam; PF03265; DNase_II; 1.
CS	DR Hypothetical protein; Hydrolase; Signal; Alternative splicing.
CT	SIGNAL 1 17 POTENTIAL.
CU	FT CHAIN 18 378 HYPOTHETICAL PROTEIN K04H4.6.
CV	FT CARBOHYD 51 51 N-LINKED (GLCNAC. . . ) (POTENTIAL).
CW	FT CARBOHYD 92 92 N-LINKED (GLCNAC. . . ) (POTENTIAL).

Submitted  
[3]  
SEQUENCE ON

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Krieser R.J., Eastman A.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Hydrolyzes DNA under acidic conditions with a preference
CC for double-stranded DNA. Has a possible role in apoptosis.
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to nucleoside 3'-
CC phosphates and 3'-phosphonucleoside end-products.
CC -!- SUBCELLULAR LOCATION: Lysosomal.
CC -!- SIMILARITY: Belongs to the DNase II family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF047017; AAC77367.1; -.
CC InterPro; IPR004947; DNase_II.
CC Pfam; PF03265; DNase_II; 1.
KW Hydrolase; Endonuclease; Nuclease; Lysosome; Zymogen; Apoptosis.
FT NON TER 1
FT PROPEP 21 20 DEOXYRIBONUCLEASE II.
FT CHAIN 21 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 276 AA; 30528 MW; 2982CBEL194B4C2B CRC64;

Query Match 23.7%; Score 459; DB 1; Length 276;
Best Local Similarity 35.8%; Pred. No. 2.9e-32;
Matches 95; Conservative 52; Mismatches 108; Indels 10; Gaps 5;

QY 98 SNNATLYVNDGVPKPVNY--SRKYGHTKGLLWNRVQGFLLHSHIPQFPP-IPREGYD 154
DB 1 NSQLAFVLYNDQPKSEKSSRGHTKGLVLLDQGGFLLHSHVNPFPASSAAYSW 60
QY 155 PPTGRNGQSGICITFKYNYEIDSOLLVNCNVYSCSIPATFHQLHMPQLCTPASS 214
DB 61 PPGAQYGTLLICVSPFLTQFLDISKQLYTYPLVYDHRLEGDFGQKFPYLEEVKGHHV 120
QY 215 SEIPGRLTTLQSAQOKELHFAKSDSFLDDIFAANMAQRLKTHLLTETWORKRQELPSN 274
DB 121 RQGFVNSSVTLTKKGFATFOSFAKFGNFGDGLYSGLWLAELGSTLQVQFQWRSGLILPSN 180
QY 275 CSLPYHYVNI--KALKLSRHSYFSSYQDHAKWCISQKTKNRWTCIGDLNRSHPQAFPSG 332
DB 181 CSGAQHVFDVDTQTAFFPAGPAGFAFNATEDHSCWCTPKGP--MACVGMNMRNQREHRGG 237
QY 333 GFICTQNWQIYQAPQGLVLYESCK 357
DB 238 GTLCAQ--MLWAKFKPLVKAWEPC 260

RESULT 8
NEUL HUMAN STANDARD; PRT; 704 AA.
AC O9BYT8; Q9ULJ4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurolysin, mitochondrial precursor (EC 3.4.24.16) (Neurotensin
DE endopeptidase) (Mitochondrial oligopeptidase M) (Microsomal
DE endopeptidase) (MEP).
GN NLN OR KIAA1226.
OS Homo sapiens (Human).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Chen J.M., Rawlings N.D., Barrett A.J.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Hydrolyzes oligopeptides such as neurotensin,
CC bradykinin, dynorphin A, etc. (By similarity).
CC -!- CATALYTIC ACTIVITY: Preferential cleavage in neurotensin: 10-
CC Pro-|-Tyr-Ile.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Mitochondrial intermembrane space and
CC also cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to peptidase family M3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ300837; CAC27329.1; -.
CC InterPro; IPR006025; Pept_M_Zn_BS.
CC InterPro; IPR001567; Peptidase_M3.
CC Pfam; PF01432; Peptidase_M3; 1.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Metalloprotease; Hydrolase; Zinc; Mitochondrion; Transit peptide.
FT TRANSIT 1 37 MITOCHONDRION (BY SIMILARITY).
FT CHAIN 38 704 NEUROLYSIN.
FT METAL 497 497 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 498 498 BY SIMILARITY.
FT METAL 501 501 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 504 504 ZINC (CATALYTIC) (BY SIMILARITY).
SQ SEQUENCE 704 AA; 80651 MW; 80136688D79BBEDF CRC64;

Query Match 4.9%; Score 95; DB 1; Length 704;
Best Local Similarity 18.7%; Pred. No. 2.4;
Matches 61; Conservative 49; Mismatches 117; Indels 100; Gaps 14;

QY 20 VLGAATISCRNE---EGKAVDMFTFYKLPKRONKE---SGETGLVLYLDSTRSRWS 72
DB 337 ILNLKKECKDRGFYDQKINADLYVYVQTEELKSIDQELFKYFFIEVVT----- 390
QY 73 EQLMNDTKSVLGRTLQOLYEAAYASKNNYALI-----YNDGVPK--PVNYSR 118
DB 391 EGLINTYQELLGLSFEQMTDAHVWNKSVTLTYVKDKATGEVLGFYLDLVPREGKYNHAA 450
QY 119 KYGHTKGLLWNRVQGFLLHSHIPQFPIPEGVDYPTGTRNQSGSGICITFKYNYEAI 178
DB 451 CFGLOPGCLL-----PDGSEMMVAALVWNF----- 476
QY 179 DSQLLVNCNVYSCSIPATFHQLH-MPQLCTPASS-----EIPRLIT- 224
DB 477 -SQPVAGRPSLLRHDVRYTFEFGVHWHQICACTDFARFSGTNTVETDFVEVPQMLNW 535
QY 225 LQSAQGGKFL--HFAKSDSFLDDIFAANMAQRL-KTHLLT---ETWQKRGELPNSCLSP 278

```

Db 536 VMDVDSLRSLSKHYKDGSPFIADDLLEKLKVASRLVNTGLTLRQIVLSKVDQSLHTNTSL- 594  
QY 279 YHYVNIKAIKUSRHYSFYSDQHPAKWC 305  
Db 595 -----DAASEYAKYC 604

RESULT 9

ID	PSUL_YEAST	STANDARD;	PRT;	970 AA.
AC	PSJ3550;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	mRNA decapping protein 2 (PSUL protein).			
CN	DCP2 OR PSUL OR INL18C OR N1917.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.			
NCBI_TaxID=4932;				
[1]				
RN	SEQUENCE FROM N.A.			
RP	SEQUENCE FROM N.A.			
EC	STRAIN=D273-10B;			
RA	Izsgoloff A.A.;			
RL	Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=97245296; PubMed=9090055;			
RA	de Antoni A., D'Angelo M.; Dal Pero F., Sartorello F., Pandolfo D.,			
RA	Pallavicini A., Lanfranchi G., Valle G.;			
RT	"The DNA sequence of cosmid 14-13b from chromosome XIV of			
RT	Saccharomyces cerevisiae reveals an unusually high number of			
RT	overlapping open reading frames.";			
RL	Yeast 13:261-266(1997).			
RN	[3]			
RP	INTERACTION WITH DCP1.			
RX	MEDLINE=99438017; PubMed=10508173;			
RA	Dunkley T., Parker R.;			
RT	"The DCP2 protein is required for mRNA decapping in Saccharomyces			
RT	cerevisiae and contains a functional Mutr motif.";			
EMBO J.	18:5411-5422(1999).			
-I-	FUNCTION: REQUIRED FOR THE PRODUCTION OF ACTIVE DECAPPING ENZYME,			
CC	PERHAPS IN A PROCESS REQUIRING THE HYDROLYSIS OF A PYROPHOSPHATE			
CC	BOND. DECAPPING IS A THE MAJOR PATHWAY OF MRNA DEGRADATION IN			
CC	YEAST. IT OCCURS THROUGH DEADENYLATION, DECAPPING AND SUBSEQUENT			
CC	5' TO 3' EXONUCLEOLYTIC DECAY OF THE TRANSCRIPT BODY.			
-I-	SUBUNIT: INTERACTS WITH DCP1.			
CC	-I- SIMILARITY: BELONGS TO THE NUDIX HYDROLASE FAMILY. STRONG, TO			
CC	S. POMBE SPAC19A8.12.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
EMBL;	L43065; AAA68866.1; -			
DR	EMBL; Z69382; CAA93389.1; -			
DR	EMBL; Z71394; CAA95998.1; -			
DR	PIR; S63059; S63059.			
DR	GeneOnline; 145124; -			
DR	SGD; S0005062; DCP2.			
DR	GO; GO:0009032; C:cytoplasmic mRNA processing body; IDA.			
DR	GO; GO:0005634; C:nucleus; IDA.			
DR	GO; GO:0008047; F:enzyme activator activity; IDA.			
DR	GO; GO:0003729; F:mRNA binding; IPI.			
DR	GO; GO:0000290; P:deadenylation-dependent decapping; IDA.			
DR	InterPro; IPR007722; DCP2.			
DR	InterPro; IPR000086; NUDIX_hydrolase.			
DR	Pfam; PF05026; DCP2; 1.			

Fri Oct 15 10:01:08 2004

us-10-790-589-4.rsp

DR InterPro; IPR001930; Peptidase M1.  
 DR Pfam; PF01433; Peptidase M1; 1.  
 DR PRINTS; PR00756; ALADIPTRASE.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 KW Hydrolase; Metalloprotease; Aminopeptidase; Zinc; Complete proteome.  
 FT METAL 272 273 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT SITE 273 273 BY SIMILARITY.  
 FT METAL 276 276 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 295 295 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT SITE 358 358 PROTON DONOR (POTENTIAL).  
 SQ SEQUENCE 781 AA; 90370 MW; 58BBF8386DF623FB CRC64;  
 Query Match 4.8%; Score 93.5; DB 1; Length 781;  
 Best Local Similarity 21.2%; Pred. No. 3.7;  
 Matches 67; Conservative 42; Mismatches 114; Indels 93; Gaps 14;  
 QY 38 WFTFYKLPKQKESGETGLEVLVLDST-----RSWRKSEQLMMDTKSVLGRITQ 88  
 Db 316 WGTFI-----NSETSGA-----LFRSLTTHPIEAHVTSPEIEQLFDDISVGKGASIL 365  
 QY 89 QLYEAVASKNNNTAYLIYNDGPKVYNYGRKYGHKGLLLMRVQ----- 133  
 Db 366 RMEAVLGDED-----FRKGIQYLN-TVKYSNATGSDFWNSLEKSGKPVSEIVKDWI 418  
 QY 134 ----GFWLIHSIOPPIPEG-----YDYPTGRNGSGGICITPKYQYEA 177  
 Db 419 TKDGYPVVYVYVNGSKINLEQERFYLKNGKNVAVYVPLTFLVNGRIYVLLKEKDSID 478  
 QY 178 IDSQLLVCNPN-----VYSCSIPATPHQELIHPQ--LCTRASSEIPGRL-LTTIQ 226  
 Db 479 IGSIDIKSVKINIDRTGFYRYVNDLSVFNKSLHDKWGLFNDYFNFFLAGRVNVTYE 538  
 QY 227 SAQQKFLHFAKSDSFL-----DDIPAAWAQRLKTHLTFTWQKQRLPNCISLPYH 280  
 Db 539 STIAQK-----FMKDDNYLVDELVSLEYLWRVNRDKYKLYEV-----LPYQ 581  
 QY 281 V--YNIKAIKLSHSY 294  
 Db 582 VKRFSKREKDELSRRTY 597  
 RESULT 11  
 NEBU HUMAN STANDARD; PRT; 6669 AA.  
 AC P20929; Q15346;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Nebulin.  
 GN NEB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95257391; PubMed=7739042;  
 RA Labeit S., Kolmerer B.;  
 RT "The complete primary structure of human nebulin and its correlation  
 to muscle structure.";  
 RL J. Mol. Biol. 248:308-315(1995).  
 RN [2]  
 RP PARTIAL PRELIMINARY SEQUENCE.  
 RX MEDLINE=88284704; PubMed=3397062;  
 RA Zeviani M., Darras B.T., Rizzuto R., Salviati G., Betto R.,  
 RA Bonilla E., Miranda A.F., Du J., Samitt C., Dickson G., Walsh F.S.,  
 RA Dinauro S., Francke U., Schon E.A.;  
 RT "Cloning and expression of human nebulin cDNAs and assignment of the  
 gene to chromosome 2q31-q32.";  
 RL Genomics 2:249-256(1988).  
 RN [3]  
 RP STRUCTURE BY NMR OF 6610-6669.  
 RX MEDLINE=98179559; PubMed=9514727;  
 RA Politou A.S., Millevoi S., Gautel M., Kolmerer B., Pastore A.;  
 "SH3 in muscles: solution structure of the SH3 domain from nebulin.";  
 J. Mol. Biol. 276:189-202(1998).  
 CC FUNCTION: This giant muscle protein may be involved in maintaining  
 the structural integrity of sarcomeres and the membrane system  
 associated with the myofibrils. Bind and stabilize F-actin.  
 CC TISSUE SPECIFICITY: Muscle specific. Located in the thin filament  
 of striated muscle.  
 CC DISEASE: Defects in NEB are a cause of the autosomal recessive  
 form of nemaline myopathy (NEM2).  
 CC SIMILARITY: Contains 1 SH3 domain.  
 CC SIMILARITY: Contains 178 nebulin repeats.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 or send an email to license@isb-sib.ch).  
 EMBL; X83957; CAA58788.1; --  
 EMBL; M19668; AAS59916.1; ALT\_SEQ.  
 EMBL; M19669; AAS59917.1; ALT\_SEQ.  
 PIR; A29979; A29979.  
 PIR; B29979; B29979.  
 PIR; S55024; S55024.  
 PDB; 1ARK; 28-JAN-98.  
 PDB; 1NEB; 24-DEC-97.  
 Genew; HGNC:7720; NEB.  
 MIM; 161650; --  
 MIM; 256030; --  
 GO; GO:0015629; C:actin cytoskeleton; TAS.  
 GO; GO:0030017; C:sarcomere; NAS.  
 GO; GO:0008307; F:structural constituent of muscle; TAS.  
 GO; GO:0030832; P:regulation of actin filament length; NAS.  
 GO; GO:0007525; P:somatic muscle development; NAS.  
 InterPro; IPR001452; SH3.  
 Pfam; PF00880; Nebulin; 146.  
 Pfam; PF00018; SH3; 1.  
 PRINTS; PR00510; NEBULIN.  
 PRINTS; PR00452; SH3DOMAIN.  
 PRODOM; PD000066; SH3; 1.  
 PROSITE; PS50002; SH3; 1.  
 Actin-binding; Muscle protein; Cytoskeleton; Repeat; 3D-structure.  
 REPEAT 76 107 NEBULIN 1.  
 REPEAT 113 143 NEBULIN 2.  
 REPEAT 148 178 NEBULIN 3.  
 REPEAT 183 213 NEBULIN 4.  
 REPEAT 218 248 NEBULIN 5.  
 REPEAT 253 283 NEBULIN 6.  
 REPEAT 289 318 NEBULIN 7.  
 REPEAT 324 354 NEBULIN 8.  
 REPEAT 363 393 NEBULIN 9.  
 REPEAT 398 428 NEBULIN 10.  
 REPEAT 434 464 NEBULIN 11.  
 REPEAT 502 532 NEBULIN 12.  
 REPEAT 537 567 NEBULIN 13.  
 REPEAT 573 603 NEBULIN 14.  
 REPEAT 611 641 NEBULIN 15.  
 REPEAT 681 711 NEBULIN 16.  
 REPEAT 749 779 NEBULIN 17.  
 REPEAT 784 814 NEBULIN 18.  
 REPEAT 820 850 NEBULIN 19.  
 REPEAT 858 888 NEBULIN 20.  
 REPEAT 893 923 NEBULIN 21.  
 REPEAT 924 954 NEBULIN 22.  
 REPEAT 959 990 NEBULIN 23.  
 REPEAT 993 1023 NEBULIN 24.  
 REPEAT 1028 1058 NEBULIN 25.  
 REPEAT 1064 1094 NEBULIN 26.  
 REPEAT 1102 1132 NEBULIN 27.





DE DE Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) (G6PD)  
 GN ZW OR G6PD OR CG12529.  
 OS Drosophila melanogaster (Fruit fly).  
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 CC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88255872; PubMed=2833831;  
 RA Fouts D.; Ganguly R.; Gutierrez A.G.; Lucchesi J.C.; Manning J.E.;  
 RT "Nucleotide sequence of the Drosophila glucose-6-phosphate  
 RT dehydrogenase gene and comparison with the homologous human gene.";  
 RL Gene 63:261-275(1988).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=P23.3, P24.1, MT41, MT68, Z3, Z5, Z11, Z16, Z21, Z27, Z41, Z42,  
 RC Z55, Z62, Z64, and Z74;  
 RX MEDLINE=97070821; PubMed=8913747;  
 RA Banes W.F.; Kirchner M.; Yoon J.; Biermann C.H.; Wang I.N.;  
 RA McCartney M.A.; Verrelli B.C.;  
 RT "Historical selection, amino acid polymorphism and lineage-specific  
 RT divergence at the G6pd locus in Drosophila melanogaster and D.  
 RT simulans.";  
 RL Genetics 144:1027-1041(1996).  
 [3]  
 RP SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).  
 RC STRAIN=Berkely;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D.; Celniker S.E.; Holt R.A.; Evans C.A.; Gocayne J.D.;  
 RA Amanatides P.G.; Scher S.E.; Li P.W.; Hoskins R.A.; Galie R.F.;  
 RA George R.A.; Lewis S.E.; Richards S.; Ashburner M.; Henderson S.N.;  
 RA Sutton G.G.; Wortman J.R.; Yandell M.D.; Zhang Q.; Chen L.X.;  
 RA Brannon R.C.; Rogers Y.-H.C.; Blazek R.G.; Champe M.; Pfeiffer B.D.;  
 RA Wan K.H.; Doyle C.; Baxter E.G.; Helt G.; Nelson C.R.; Miklos G.L.G.;  
 RA Abril J.F.; Agbayani A.; An H.-J.; Andrews-Pfannkoch C.; Baldwin D.;  
 RA Ballew R.M.; Basu A.; Baxendale J.; Bayraktaroglu L.; Beasly E.M.;  
 RA Beeson K.Y.; Benos P.V.; Berman B.P.; Brokstein P.; Brotter P.;  
 RA Borkova D.; Botchan M.R.; Bouck J.; Brockstein P.; Cantor A.; Chandra I.;  
 RA Burtis K.C.; Busam D.A.; Butler H.; Cadieu E.; Center A.; Chandra I.;  
 RA Cherry J.M.; Cawley S.; Dahlke S.; Davenport L.B.; Davies P.;  
 RA de Pablo B.; Delcher A.; Deng Z.; Mays A.D.; Dew I.; Dietz S.M.;  
 RA Dodson K.; Doup L.B.; Downes M.; Dugan-Rocha S.; Dunkov B.C.; Dunn P.;  
 RA Durbin K.J.; Evangelista C.C.; Ferraz C.; Ferreira S.; Fleischmann W.;  
 RA Fosler C.; Gabrielian A.E.; Garg N.S.; Gelbart W.M.; Glasser K.;  
 RA Glodek A.; Gong F.; Gorrell J.H.; Gu Z.; Guan P.; Harris M.;  
 RA Harris N.L.; Harvey D.A.; Heiman T.J.; Hernandez J.R.; Houck J.;  
 RA Hostin D.; Houston K.A.; Howland T.J.; Wei M.-H.; Ibegwam C.;  
 RA Jalaal M.; Kalush F.; Karpen G.H.; Ke Z.; Kennison J.A.; Ketchum K.A.;  
 RA Kimmel B.E.; Kodira C.D.; Kraft C.; Kravitz S.; Kulp D.; Lai Z.;  
 RA Lasko P.; Lei Y.; Levitsky A.A.; Li J.H.; Li Z.; Liang Y.; Lin X.;  
 RA Liu X.; Mattei B.; McIntosh T.C.; McLeod M.P.; McPherson D.;  
 RA Merkulov G.; Milshina N.V.; Mobarry C.; Morris J.; Moshrefi A.;  
 RA Mount S.M.; Moy M.; Murphy B.; Murphy L.; Muzny D.M.; Nelson D.L.;  
 RA Nelson D.R.; Nelson K.A.; Nixon K.; Nuskern D.R.; Pacle J.M.;  
 RA Palazzolo M.; Pittman G.S.; Pan S.; Pollard J.; Puri V.; Reese M.G.;  
 RA Reinert K.; Remington K.; Saunders R.D.C.; Scheeler F.; Shen H.;  
 RA Shue B.C.; Siden-Kiamos I.; Simpson M.; Skupski M.P.; Smith T.;  
 RA Spier E.; Spradling A.C.; Stapleton M.; Strong R.; Sun E.;  
 RA Svirskas R.; Tector C.; Turner R.; Venter E.; Wang A.H.; Wang X.;  
 RA Wang Z.-Y.; Wasserman D.A.; Weinstein G.M.; Weissbach J.;  
 RA Williams S.M.; Woodage T.; Worley K.C.; Wu D.; Yang S.; Yao Q.A.;  
 RA Ye J.; Yeh R.-F.; Zaveri J.S.; Zhan M.; Zhang Q.; Zheng L.;  
 RA Zheng X.H.; Zhong F.N.; Zhong W.; Zhou X.; Zhu S.; Zhu X.; Smith H.O.;  
 RA Gibbs R.A.; Myers E.W.; Rubin G.M.; Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 [4]  
 RN SEQUENCE FROM N.A. (ISOFORM LONG).  
 RP STRAIN=Berkely; TISSUE=Embryo;  
 RC MEDLINE=22426066; PubMed=12537569;  
 RX Stapleton M.; Carlson J.W.; Brokstein P.; Yu C.; Champe M.,

George R.A., Guarin H., Krommiller B., Pacle J.M., Park S., Wan K.H.,  
 Rubin G.M., Celniker S.E.;  
 "A Drosophila full-length cDNA resource.";  
 Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).  
 -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate + NADP(+) = D-glucono-  
 1,5-lactone 6-phosphate + NADPH.  
 -!- PATHWAY: Pentose phosphate pathway; first step.  
 -!- ALTERNATIVE PRODUCTS:  
 Event=Alternative splicing; Named isoforms=2;  
 Name=Long;  
 IsoId=Pi2646-1; Sequence=Displayed;  
 Name=Short;  
 IsoId=Pi2646-2; Sequence=VSP\_001593;  
 Note=No experimental confirmation available;  
 -!- SIMILARITY: Belongs to the glucose-6-phosphate dehydrogenase  
 family.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 or send an email to license@isb-sib.ch).  
 -----  
 EMBL; M26674; AAA51463.1; -;  
 EMBL; M26673; AAA51463.1; JOINED.  
 EMBL; U42738; AAB02801.1; -;  
 EMBL; U42739; AAB02802.1; -;  
 EMBL; U42740; AAB02803.1; -;  
 EMBL; U42741; AAB02804.1; -;  
 EMBL; U42742; AAB02805.1; -;  
 EMBL; U42743; AAB02806.1; -;  
 EMBL; U42744; AAB02807.1; -;  
 EMBL; U42745; AAB02808.1; -;  
 EMBL; U42746; AAB02809.1; -;  
 EMBL; U42747; AAB02810.1; -;  
 EMBL; U42748; AAB02811.1; -;  
 EMBL; U42749; AAB02812.1; -;  
 EMBL; U43165; AAA99071.1; -;  
 EMBL; U43166; AAA99072.1; -;  
 EMBL; U43167; AAA99073.1; -;  
 EMBL; U44721; AAA99092.1; -;  
 EMBL; U44722; AAA99107.1; -;  
 EMBL; AE003512; AAF48999.1; -;  
 EMBL; AE003512; AAF49000.1; -;  
 EMBL; AY052079; AAK93503.1; -;  
 PIR; A47740; A47740.  
 PIR; J02072; DEFPG6.  
 HSP; P11411; IDPG.  
 FlyBase; FBgn004057; Zw.  
 InterPro; IPR001282; G6PD.  
 Pfam; PF00479; G6PD; 1.  
 Pfam; PF02781; G6PD; C; 1.  
 PRINTS; PR00079; G6PDHGRNASE.  
 ProDom; PD001129; G6PD; 1.  
 TIGRFAMs; TIGR00871; zwf; 1.  
 PROSITE; PS00069; G6P DEHYDROGENASE; 1.  
 Oxioreductase; NADP; Glucose metabolism; Polymorphism;  
 Alternative splicing.  
 ACT SITE 209 209  
 VARSPLIC 1 7  
 BY SIMILARITY.  
 Missing (in isoform Short).  
 /FTID=VSP\_001593.  
 G -> C (IN STRAINS P24.1, MT32 AND MT68).  
 T -> N (IN STRAIN Z74).  
 L -> P (IN STRAINS P23.3, MT41, Z3, Z5,  
 Z11, Z16, Z21, Z27, Z42, Z55, Z64, Z74  
 AND BERKELEY).  
 QA -> AG (IN REF. 1).  
 N -> K (IN REF. 1).  
 LGV -> ARS (IN REF. 1).  
 DELRE -> AAAQ (IN REF. 1).  
 SEQUENCE 524 AA; 60431 MW; ABF81B763A82F1FD CRC64;

Query Match 4.8%; Score 92; DB 1; Length 524;  
 Best Local Similarity 18.3%; Pred. No. 2.9;  
 Matches 50; Conservative 41; Mismatches 88; Indels 94; Gaps 12;

QY 101 TAVLIYNDG-VKPVNYSKYGHGKGLLNVRVQGFWLHISIQFPPPEEGYDYP--- 156  
 Db 55 TLWMLYRDDLLPKPTKFC---GYARGLTVDSIKE---QCLPYMKVQPHQKQYEFWA 107  
 QY 157 -----TGRNGSGICITFKINOYEALDSOLLY-----CNPVYSCSIPATFHOBLIHM 205  
 Db 108 LNEYVSGYDGRG-----FELLNQLEIMENKANKRIFYLALPPSFERYVTNMI 158  
 QY 206 POLCTR-----ASSEIPGRLLTLTQSAQGFHLHFAKSDFLDDIFA 248  
 Db 159 KQICMSVCGWNRVIRKPPGRDDASSQALSHLAGLFQEDQLYRIDHYLGE----- 210  
 QY 249 ANMAQRLKT-----HLITITWQKQELPSNLSPLHVYVNIKAUKS-----RH5Y 294  
 Db 211 --MVQNLMITIRGNKILSTWNR-----NIASVLITFKPEFTQGRGGY 253  
 QY 295 FGSY-----QDHAKWCISQKGTNRWTCIGD 320  
 Db 254 FDEFGIIRDMVQNHLLQILSLVAMEKPVSCHPD 286

## RESULT 13

ID HAPI HAEIN STANDARD; PRT; 1409 AA.  
 AC P44596;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Adhesion and penetration protein precursor (BC 3.4.21.-).  
 GN HAP OR HI0248.  
 OS Haemophilus influenzae.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 OC Pasteurellaceae; Haemophilus.  
 OX NCBI\_TaxID=727;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=rd / KW20 / ATCC 51907;  
 RX MEDLINE=95350630; PubMed=7542800;  
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 RA McKerney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae  
 Rd.";  
 RL Science 269:496-512(1995).  
 CC -!- FUNCTION: PROBABLE PROTEASE; PROMOTES ADHERENCE AND INVASION BY  
 CC DIRECTLY BINDING TO A HOST CELL STRUCTURE (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Secreted (Potential).  
 CC -!- DOMAIN: THE SIGNAL PEPTIDE GUIDE THE PRECURSOR TO THE PERIPLASMIC  
 CC SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE  
 CC OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE  
 CC DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY  
 CC SIMILARITY).  
 CC -!- SIMILARITY: Belongs to peptidase family S6.  
 CC -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A STOP CODON HAD TO  
 CC BE SKIPPED IN POSITION 710 TO PRODUCE THIS ORF.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 -----  
 CC EMBL; U32710; -; NOT\_ANNOTATED\_CDS.  
 CC MEROPS; S06.006; -.  
 DR TIGR; HI0248; -.  
 DR InterPro; IPR006315; Autotransport.  
 DR InterPro; IPR005546; Autotransporter.  
 DR InterPro; IPR000710; Peptidase S6.  
 DR InterPro; IPR004899; Peptactin.  
 DR Pfam; PF02395; IGAL1; 1.  
 DR Pfam; PF03212; Peptactin; 1.  
 DR PRINTS; PR00921; IGASERPTASE.  
 DR TIGR; TIGR01414; autotrans\_bar1; 1.  
 KW Hydrolyase; Serine protease; Transmembrane; Zymogen; Signal;  
 KW Complete proteome.  
 FT SIGNAL 1 25 POTENTIAL.  
 FT CHAIN 26 ? ADHESION AND PENETRATION PROTEIN.  
 FT PROPEP ? 1409 HELPER PEPTIDE (POTENTIAL).  
 FT ACT\_SITE 250 250 BY SIMILARITY.  
 FT ACT\_SITE 250 250 BY SIMILARITY.  
 SQ SEQUENCE 1409 AA; 156797 MW; 63ABC893FA84D16E CRC64;  
 Query Match 4.8%; Score 92; DB 1; Length 1409;  
 Best Local Similarity 20.4%; Pred. No. 11;  
 Matches 76; Conservative 49; Mismatches 124; Indels 124; Gaps 19;

QY 32 EGAVDWFTF-YKLPKRNKESGETGLEY-----LYL 62  
 Db 110 EGRNPDQHRFTYQIVKRNYYQAWERKHPYDGDYHMPRLHKFVTEAEFVGMVTNMDGKYA 169  
 QY 63 D-----STTSRWKSEQLMNDTKS-----VLGRTLOOLYEAAYASKNN 100  
 Db 170 DRENPERVRIGSGQYWRDKDBETNVHSSYYVSGAYRYLTAGNTHQ-----SCNGNG 224  
 QY 101 TAYLIYNDGVKPVNYS-RKYGHTKG-----LLLNVRVQGFWLHISIQFPPPEEGYD 154  
 Db 225 TVNLSEN--VVSFNHYGPLPTGSGKSGSPMFYIDAKKQWLNAVILQ-----TGH 274  
 QY 155 PPTGRNGSGICITFKINOYEALDSOLLVCPNVYSCSI-PATFHOBLIHMPCQTRAS 213  
 Db 275 PFGRGNGFQLIREBEWFYNEVLAVDT-----PSVFQRYIPIPNHYSFV-----S 319  
 QY 214 SSEIPGRLLTLTQSAQGFHLHFAKSDFLDDIFAAMWAQRLKTHLLTETWQKQELPS 273  
 Db 320 NNDGTGKLTLPKPSKSGK-----AKSEVGTVKLFNPSLNTAKGVH----- 361  
 QY 274 NCSLPFHYVNIKAIKLSRHSYFSSYQDHAKWCISQKGTNRWTCIGDLNRSFHOAFRSGG 333  
 Db 362 KAAAGYNIQPR-MEYGNIVLGD-----OGKGT-----LTENNINQAGGLIPEGN 408  
 QY 334 FIC--TON---WQ 341  
 Db 409 FVWKGQNNTWQ 421

RESULT 14  
 Y348\_MYCPN STANDARD; PRT; 305 AA.  
 ID Y348\_MYCPN  
 AC P75255;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Hypothetical lipoprotein MG348 homolog precursor (Gl2\_orf305).  
 GN MEN523 OR MF319.  
 OS Mycoplasma pneumoniae.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2104;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 29342 / M129;  
 RX MEDLINE=97105885; PubMed=8948633;  
 RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,  
 RA Herrmann R.;

RT "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae.";  
BL Nucleic Acids Res. 24:4420-4449(1996).  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Potential).  
CC  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
CC  
CC  
CC EMBL; AB000029; AAB95967.1; --  
DR BIR; S73645; S73645.  
DR InterPro; IPR000437; Prok lipoprot S.  
DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.  
KW Hypothetical protein; Lipoprotein; Membrane; Signal; Complete proteome; Palmitate.  
FT SIGNAL 1 29 POTENTIAL.  
FT CHAIN 30 305 HYPOTHETICAL LIPOPROTEIN MG348 HOMOLOG.  
FT LIPID 30 30 N-palmitoyl cysteine (Potential).  
FT LIPID 30 30 S-diacylglycerol cysteine (Potential).  
SQ SEQUENCE 305 AA; 33536 MW; 38A143D603F4BEE3 CRC64;  
Query Match 4.7%; Score 91.5; DB 1; Length 305;  
Best Local Similarity 23.9%; Pred. No. 1.6;  
Matches 55; Conservative 33; Mismatches 75; Indels 67; Gaps 12;  
QY 7 RTSFA-LLFLGLFGVIGAA-----TISCNEEGKAVDFTFKLPKRONKE----- 51  
Db 12 KVSFSTILLGSGVILSSCNIDKPNVFTLSQSVENK-VD---YSKLPK-ENKTVRNL 66  
QY 52 -----SGENGLEYL-YLDSTTSWRKSEQLMNDTKSVLGRITLQQLYEAY 94  
Db 67 VFGTAENDQNYVLVVTETDSSQINFNGSNQAVSTENWAGD-----LGTWKQVQNY 122  
QY 95 ASKSNNTAYLIYNDGVPKPNYRKYGTGKLLMNRVQGFWLHSPQPPPIPEGYDY 154  
Db 123 STYPKGVKFLWINDIDENPKV-----WNPFRYPVIASDNLAKQTDK--DN 167  
QY 155 PPTGRRNGQGICITFKNQVHAIDSQLVCNPNVYSCSIPATPHOLIH 204  
Db 168 SDKLRNDESAI-----QYR-----EIVTFIQTVYSGVNNLINQSNVH 206  
RESULT 15  
ID ITA8 HUMAN STANDARD; PRT; 1025 AA.  
AC P53708;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Integrin alpha-8.  
GN ITGA8.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP MEDLINE=95286701; PubMed=7768999;  
RA Schnapp L.M., Breuss J.M., Ramos D.M., Sheppard D., Pytela R.;  
RT "Sequence and tissue distribution of the human integrin alpha 8 subunit: a beta 1-associated alpha subunit expressed in smooth muscle cells.";  
RT J. Cell Sci. 108:537-544(1995).  
RL J. Cell Sci. 108:537-544(1995).  
CC -!- FUNCTION: INTEGRIN ALPHA-8/BETA-1 IS A RECEPTOR FOR FIBRONECTIN AND CYTOACTIN. IT RECOGNIZES THE SEQUENCE R-G-D IN ITS LIGANDS.  
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAIN LINKED BY A DISULFIDE BOND. ALPHA-8 ASSOCIATES WITH BETA-1.  
CC

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN SMOOTH MUSCLE AND SMOOTH MUSCLE-LIKE CONTRACTILE CELLS.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
CC  
CC  
CC EMBL; L36531; AAA93514.1; --  
DR HSSP; P06756; IJW2.  
DR Genew; HGNC:6144; ITGA8.  
DR MIM; 604063; --  
DR GO; GO:0008305; C:integrin complex; TAS.  
DR GO; GO:0004895; R:cell adhesion receptor activity; NAS.  
DR GO; GO:0016337; P:cell-cell adhesion; NAS.  
DR GO; GO:0007160; P:cell-matrix adhesion; NAS.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR Pfam; PF01839; FG-GAP; 4.  
DR PRINTS; PR01185; INTEGRINA.  
DR SMART; SM00191; Int\_alpha; 6.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane; Repeat; Calcium.  
KW CHAIN 1 868 INTEGRIN ALPHA-8 HEAVY CHAIN (POTENTIAL).  
FT CHAIN 869 1025 INTEGRIN ALPHA-8 LIGHT CHAIN (POTENTIAL).  
FT DOMAIN 1 974 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 975 995 POTENTIAL.  
FT DOMAIN 996 1025 CYTOPLASMIC (POTENTIAL).  
FT REPEAT 16 79 FG-GAP 1.  
FT REPEAT 95 160 FG-GAP 2.  
FT REPEAT 161 225 FG-GAP 3.  
FT REPEAT 226 279 FG-GAP 4.  
FT REPEAT 280 345 FG-GAP 5.  
FT REPEAT 346 405 FG-GAP 6.  
FT REPEAT 409 463 FG-GAP 7.  
FT CA\_BIND 291 299 POTENTIAL.  
FT CA\_BIND 421 429 POTENTIAL.  
FT SITE 417 419 CELL ATTACHMENT SITE (POTENTIAL).  
FT FT 58 BY SIMILARITY.  
FT FT 112 BY SIMILARITY.  
FT FT 149 BY SIMILARITY.  
FT FT 469 BY SIMILARITY.  
FT FT 486 BY SIMILARITY.  
FT FT 603 BY SIMILARITY.  
FT FT 675 BY SIMILARITY.  
FT FT 829 BY SIMILARITY.  
FT FT 891 BY SIMILARITY.  
FT FT 43 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT 84 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT 139 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT 201 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT 273 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT 466 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT 563 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT 567 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT 681 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT 699 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT 715 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT 742 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT 858 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT 885 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT FT 967 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 1025 AA; 113612 MW; 58B562C00CF7B93 CRC64;  
Query Match 4.7%; Score 91; DB 1; Length 1025;



***This Page Blank (uspio)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2004, 00:33:41 ; Search time 4897.47 Seconds  
(without alignments)  
11221.906 Million cell updates/sec

Title: US-10-790-589-3  
Perfect score: 1268  
Sequence: 1 atggggaagtgctcgtg.....tctctcatgtttaccattta 1268

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Result No.	Score	Query Match %	Length	DB ID	Description
1	1268	100.0	1268	6	AR201144 Sequence
2	1140	89.9	1140	6	AR274571 Homo sapi
3	1079.6	85.1	1086	6	AR432047 Sequence
4	1079.6	85.1	1086	6	BD091715 Novel deo
5	920.6	72.6	1029	9	AF333389 Homo sapi
6	881.8	53.8	1498	10	AF178974 Rattus no
7	673.8	53.1	1652	6	AR432046 Sequence
8	673.8	53.1	1652	6	BD091714 Novel deo
9	673.8	53.1	1654	10	AF128888 Mus muscu
10	672.8	53.1	1224	6	AR201143 Sequence
11	487.8	38.5	19164	9	AF334602 Homo sapi
12	487.8	38.5	164529	2	AC027606 Homo sapi
13	487.8	38.5	205034	9	AL359273 Human DNA
14	254	20.0	2337	10	AF334603S6
15	254	20.0	197457	2	AC127571 Mus muscu
16	254	20.0	211003	2	AC141632 Mus muscu
17	254	20.0	255925	2	AC114618 Mus muscu
18	246.4	19.4	231698	2	AC118117 Rattus no
19	246.4	19.4	306870	2	AC098557 Rattus no
20	179	14.1	86816	2	AL136113 Homo sapi
21	160	12.6	1357	9	AB074093 Homo sapi
22	158.4	12.5	1387	9	AB074094 Pan trogl
23	156.8	12.4	1388	9	AB074095 Gorilla g
24	156.4	12.3	1388	9	AB074096 Pongo pyg
25	155.2	12.2	1390	9	AB074097 Hyllobates
26	155.2	12.2	1391	9	AB074099 Macaca mu
27	155	12.2	7470	14	F066408 Fowlpox v
28	155	12.2	288539	14	AF198100 Fowlpox v
29	153.6	12.1	1388	9	AB074098 Papio han
30	153.6	12.1	1390	9	AB074100 Macaca fa
31	148.8	11.7	1391	9	AB074101 Aotus tri
32	114	9.0	642	10	AF334603S5
33	97.4	7.7	1374	10	AF334603S2
34	89.2	7.0	719	10	AF334603S4
35	84.8	6.7	1724	9	AK075449
36	84	6.6	1593	9	AK004574
37	84	6.6	1920	9	AF047016
38	82.4	6.5	1083	9	BT007047
39	82.4	6.5	1083	12	BT007514
40	82.4	6.5	1575	6	BC337903
41	82.4	6.5	1601	9	BC010419
42	82.4	6.5	1751	9	AF045937
43	82.4	6.5	1825	10	AF334603S1
44	82.4	6.5	1975	9	AF060222
45	73	5.8	1292	4	S6J001387

# ALIGNMENTS

RESULT 1	AR201144	Sequence 3	1268 bp	DNA	linear	PAT 20-APR-2002
LOCUS	AR201144	Sequence 3	from patent US 6358723.			
DEFINITION	AR201144					
ACCESSION	AR201144					
VERSION	AR201144.1	GI:20252032				
KEYWORDS						
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 1268)					
AUTHORS	Eastman, A. Richard, and Krieser, R. Joe.					
TITLE	Deoxyribonuclease II beta, proteins and cDNAs					
JOURNAL	Patent: US 6358723-A 3 19-MAR-2002;					
FEATURES	Location/Qualifiers					

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

[illegible]



Query Match									
Best Local Similarity 89.9%; Score 1140; DB 9; Length 1140;									
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	ATGGGGAAGTCTCTGCTGTGGCATGAAATAAATGAACACAAAAATGATGCCAAGCTG	60						
DB	1	ATGGGGAAGTGTCTGCTGTGGCATGAAATAAATGAACACAAAAATGATGCCAAGCTG	60						
QY	61	CTAAGAACATCCTTTGCTTTGCTCTTCTCTTGCCCTCTTTGGGCGTCTGGGGGACGACA	120						
DB	61	CTAAGAACATCCTTTGCTTTGCTCTTCTCTTGCCCTCTTTGGGCGTCTGGGGGACGACA	120						
QY	121	ATTTTCATCGAAGATGAAGAGGGAAGAGCTGGACTGGTTTACTTTTTATAAGTTA	180						
DB	121	ATTTTCATCGAAGATGAAGAGGGAAGAGCTGGACTGGTTTACTTTTTATAAGTTA	180						
QY	181	AAAAGACAAAAAAGGAAAGTGGAGAGCTGGTTAGAGTACCTGTACTAGACTCTACA	240						
DB	181	AAAAGACAAAAAAGGAAAGTGGAGAGCTGGTTAGAGTACCTGTACTAGACTCTACA	240						
QY	241	ACTAGAAGCTGGAGGAAGAGTGAGCAACTAATGAATGACACCAAGAGTGTTTGGGAAG	300						
DB	241	ACTAGAAGCTGGAGGAAGAGTGAGCAACTAATGAATGACACCAAGAGTGTTTGGGAAG	300						
QY	301	ACATTTACACAGCTATATGAAGCATATGCCCTTAAGAGTAAACAACAGACCTATCTAATA	360						
DB	301	ACATTTACACAGCTATATGAAGCATATGCCCTTAAGAGTAAACAACAGACCTATCTAATA	360						
QY	361	TACAATGATGAGTCCCTTAAACCTGTGAAATACAGTAGAAGATGAGCACACCAAGGT	420						
DB	361	TACAATGATGAGTCCCTTAAACCTGTGAAATACAGTAGAAGATGAGCACACCAAGGT	420						
QY	421	TTACTGCTGTGGAACAGAGTTTCAAGGGTCTTGGGTGATTCAATCCATCCCTCAGTTTCT	480						
DB	421	TTACTGCTGTGGAACAGAGTTTCAAGGGTCTTGGGTGATTCAATCCATCCCTCAGTTTCT	480						
QY	481	CCAAATTCGGGAAGAGGCTATGATATCCACCCACAGGAGACGAAATGGAACAAAGTGGC	540						
DB	481	CCAAATTCGGGAAGAGGCTATGATATCCACCCACAGGAGACGAAATGGAACAAAGTGGC	540						
QY	541	ATCTGCATAACTTTCAAGTACAAACAGTATGAGCGAATAGATTCTCAGCTCTTGGTCTGC	600						
DB	541	ATCTGCATAACTTTCAAGTACAAACAGTATGAGCGAATAGATTCTCAGCTCTTGGTCTGC	600						
QY	601	AAOCCCAACGTTCTAGCTGCTCCATCCAGGCCACTTTCACAGGAGCTCAATTCACATG	660						
DB	601	AAOCCCAACGTTCTAGCTGCTCCATCCAGGCCACTTTCACAGGAGCTCAATTCACATG	660						
QY	661	CCCCAGCTGTGCCACAGGGCCAGCTCATCAGAGATTCCTGGCAGGCTCCTCACACACTT	720						
DB	661	CCCCAGCTGTGCCACAGGGCCAGCTCATCAGAGATTCCTGGCAGGCTCCTCACACACTT	720						
QY	721	CAGTCGGCCACAGGACAAAAATTCCTCCATTTTGCAGAGTCGGATTCCTTTCTTGA	780						
DB	721	CAGTCGGCCACAGGACAAAAATTCCTCCATTTTGCAGAGTCGGATTCCTTTCTTGA	780						
QY	781	ATCTTTCAGCCTGGATGGTCTCAACGGCTGAAGACACACTTGTGTAAACAGAAACCTGGCAG	840						
DB	781	ATCTTTCAGCCTGGATGGTCTCAACGGCTGAAGACACACTTGTGTAAACAGAAACCTGGCAG	840						
QY	841	CGAAAAAGACAGAGCTTCTTCAAACTGCTCCCTTCCCTTACCATGCTCAATATAAAA	900						
DB	841	CGAAAAAGACAGAGCTTCTTCAAACTGCTCCCTTCAAACTGCTCAATATAAAA	900						
QY	901	GCAATTAATTTATCAGACACTCTTATTTTTCAGTTCTTATCAAGATCAGGCCAAGTGGTGT	960						
DB	901	GCAATTAATTTATCAGACACTCTTATTTTTCAGTTCTTATCAAGATCAGGCCAAGTGGTGT	960						
QY	961	ATTTTCCCAAAAGGGCACCAAAAATCGCTGACATGTATTCGAGACCTTAAATCGAGTCCA	1020						
DB	961	ATTTTCCCAAAAGGGCACCAAAAATCGCTGACATGTATTCGAGACCTTAAATCGAGTCCA	1020						
QY	1021	CACCAAGCCTTCAGAGTGGAGGATTCATTTGTACCAAGATTTGGCAAAATTTTACCAGCA	1080						

Db	1021	CACCAAGCCTTCAGAAGTGGAGGATTCATTTGTATCCACAGANTTGGCAATTTTCCACGCA	1086 bp	linear	PAT 18-DEC-2003
Qy	1081	TTTTCAGGATTTAGTATTATATCTATGAAGCTGTGAAGTAACTTGGTGAAGGACACAGGT			1140
Db	1081	TTTCAAGGATTAGTATTATCTATGAAGCTGTGAAGTAACTTGGTGAAGGACACAGGT			1140
RESULT 3					
AR432047					
LOCUS	AR432047	Sequence 4 from patent US 6653118.	1086 bp	DNA	linear
DEFINITION	AR432047	Sequence 4 from patent US 6653118.			
ACCESSION	AR432047				
VERSION	AR432047.1	GI:40194250			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1	(bases 1 to 1086)			
AUTHORS	Tanuma,S.-i. and Shikawa,D.				
TITLE	Deoxyribonuclease, gene encoding same and use thereof				
JOURNAL	Patent: US 6653118-A 4 25-NOV-2003;				
FEATURES	Location/Qualifiers				
source	1..1086				
	/organism="unknown"				
ORIGIN	/mol_type="genomic DNA"				
Query Match	85.1%;	Score 1079.6;	DB 6;	Length 1086;	
Best Local Similarity	99.6%;	Pred. No. 3.6e-291;			
Matches 1082;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;	
Qy	34	ATGAAACAGAAATGATGGCAAGACTGCTAAGACATCCCTTTCCTTGCCTTTCCTTGGC	93		
Db	1	ATGAAACAGAAATGATGGCAAGACTGCTAAGACATCCCTTTCCTTGCCTTTCCTTGGC	60		
Qy	94	CTCTTTGGGTGCTGGGGGAGCAACAAATTTTCATGCAGAAATGAAGGAAAGCTGTG	153		
Db	61	CTCTTTGGGTGCTGGGGGAGCAACAAATTTTCATGCAGAAATGAAGGAAAGCTGTG	120		
Qy	154	GACTGGTTTACTTTTATAAGTAACTTAAAGCAAAACAAGGAAAGTGAGAGACTGG	213		
Db	121	GACTGGTTTACTTTTATAAGTAACTTAAAGCAAAACAAGGAAAGTGAGAGACTGG	180		
Qy	214	TTAGAGTACCTGTACCTPAGACTCTACAACCTAGAGCTGGAGGAGAGTGCACATAATG	273		
Db	181	TTAGAGTACCTGTACCTPAGACTCTACAACCTAGAGCTGGAGGAGAGTGCACATAATG	240		
Qy	274	AATGACACCAAGAGTGTTTTGGGAAGGACATTAACAACAGCTATATGAAGCATATGCTCT	333		
Db	241	AATGACACCAAGAGTGTTTTGGGAAGGACATTAACAACAGCTATATGAAGCATATGCTCT	300		
Qy	334	AAGAGTAAACACACAGGCTATCTAATATACATGATGGAGTGCCTTAAACCTGTGAATTAC	393		
Db	301	AAGAGTAAACACACAGGCTATCTAATATACATGATGGAGTGCCTTAAACCTGTGAATTAC	360		
Qy	394	AGTAGAAGTATGGACACACAAAGTTTACTGCTGTGAAACAGAGTTCACAGGTTCTGG	453		
Db	361	AGTAGAAGTATGGACACACAAAGTTTACTGCTGTGAAACAGAGTTCACAGGTTCTGG	420		
Qy	454	CTGATTCATTCCATCCCTCAGTTTCTCCAAATTCGGGAAGAGGCTATGATTCACACC	513		
Db	421	CTGATTCATTCCATCCCTCAGTTTCTCCAAATTCGGGAAGAGGCTATGATTCACACC	480		
Qy	514	ACAGGAGACGAATGGACAAAGTGGCATCTGCATACTTTCAAGTACCAACAGTATGAG	573		
Db	481	ACAGGAGACGAATGGACAAAGTGGCATCTGCATACTTTCAAGTACCAACAGTATGAG	540		
Qy	574	GCAATAGATTCTCAGCTCTTGCTGCAACCCCAAGCTATAGCTGCTCCATCCAGCC	633		
Db	541	GCAATAGATTCTCAGCTCTTGCTGCAACCCCAAGCTATAGCTGCTCCATCCAGCC	600		
Qy	634	ACCTTTCCAGGAGCTCAATTCATGTCGCCAGCTGTGCACAGGGCCAGCTCATCAG	693		

Db	601	ACCTTTTACAGAGAGCTCATTCATGCCCCAGCTGTGCACAGGCCAGCTCATCAGAG	660
Qy	694	ATTCTGGCAGGCTCTCCACACATTCAGTCGGCCAGGGAACAAAATTCCTCCATTTT	753
Db	661	ATTCTGGCAGGCTCTCCACACATTCAGTCGGCCAGGGAACAAAATTCCTCCATTTT	720
Qy	754	GCAAGTCGGATTCCTTTCTTGACGACATCTTTTGACGCTGGATGCTCAACGGCTGAAG	813
Db	721	GCAAGTCGGATTCCTTTCTTGACGACATCTTTTGACGCTGGATGCTCAACGGCTGAAG	780
Qy	814	ACACACTTGTAAACAGAACTGGCAGCGAAAGAACAGAGCTTCCTTCAAACTGCTCC	873
Db	781	ACACACTTGTAAACAGAACTGGCAGCGAAAGAACAGAGCTTCCTTCAAACTGCTCC	840
Qy	874	CTTCTTTACCATGCTACATAATAAAGCAATTAATTAACGACACATCTTTATTTCAAGT	933
Db	841	CTTCTTTACCATGCTACATAATAAAGCAATTAATTAACGACACATCTTTATTTCAAGT	900
Qy	934	TCCTTACAGATCAGCGCAAGTGGTGTATTTCCCAAAAGGCGACCAAAATTCGCTGACA	993
Db	901	TCCTTACAGATCAGCGCAAGTGGTGTATTTCCCAAAAGGCGACCAAAATTCGCTGACA	960
Qy	994	TGATTTGGACACCTTAATCGGAGTCCACACCAAGCTTTCAGAAAGTGGAGGATTCATTTGT	1053
Db	961	TGATTTGGACACCTTAATCGGAGTCCACACCAAGCTTTCAGAAAGTGGAGGATTCATTTGT	1020
Qy	1054	ACCCAGAAATGGCAAAATTTACCAAGCAATTTCAAGGATTTAGTATTAATAGAAAGCTGT	1113
Db	1021	ACCCAGAAATGGCAAAATTTACCAAGCAATTTCAAGGATTTAGTATTAATAGAAAGCTGT	1080
Qy	1114	AAGTAA 1119	
Db	1081	AAGTAA 1086	
RESULT 4			
LOCUS	BD091715	1086 bp	DNA linear
DEFINITION	Novel deoxyribonuclease, gene encoding thereof and use thereof.		PAT 27-AUG-2002
ACCESSION	BD091715		
VERSION	BD091715.1	GI:22637326	
KEYWORDS	WO 0112793-A/2.		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL	Tanuma, S. and Shiohara, D.		
COMMENT	Novel deoxyribonuclease, gene encoding thereof and use thereof		
	Patent: WO 0112793-A 2 22-FEB-2001;		
	SEIICHI TANUMA, DAISUKE SHIOKAWA		
	OS Homo sapiens (human)		
	EN WO 0112793-A/2		
	PD 22-FEB-2001		
	PF 01-MAY-2000 WO 2000JP002893		
	PR 17-AUG-1999 JP 99P 230870		
	PI SEIICHI TANUMA, DAISUKE SHIOKAWA		
	PC C12N9/22, C12N15/12, C12N5/10, C12N7/01, A61K31/195, A61K48/00 CC		
	FT Key		
	Location/Qualifiers		
	(1)..(1083).		
FEATURES			
source	1..1086		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	85.1%;	Score 1079.6;	DB 6; Length 1086;
Best Local Similarity	99.6%;	Pred. No. 3.6e-291;	
Matches 1082;	Conservative	0; Mismatches 4;	Indels 0; Gaps 0;
Qy	34	ATGAAACAGAAATGATGCAAGACTGCTAAGAACATCCTTTGCTTTGCTTCTTCTTGGC	93

Db	1	ATGAAACAGAAATGATGCAAGACTGCTAAGAACATCCTTTGCTTTGCTTCTTCTTGGC	60
Qy	94	CTCTTTGGGCTGCTGGGGCAGCAACAATTCATGCAAGAAATCAAGAGGAAAGCTGTG	153
Db	61	CTCTTTGGGCTGCTGGGGCAGCAACAATTCATGCAAGAAATCAAGAGGAAAGCTGTG	120
Qy	154	GACTGGTTTACCTTTTATTAAGTTTACCTAAAAGACAAAACAAGAAAGTGGAGACTGGG	213
Db	121	GACTGGTTTACCTTTTATTAAGTTTACCTAAAAGACAAAACAAGAAAGTGGAGACTGGG	180
Qy	214	TTAGAGTACCTGTACCTAGACTCTCAACCTAGAAAGCTGGAGGAAAGTGGAGAACTAATG	273
Db	181	TTAGAGTACCTGTACCTAGACTCTCAACCTAGAAAGCTGGAGGAAAGTGGAGAACTAATG	240
Qy	274	AATGACACCAAGAGTGTCTTGGGAAAGCAATTAACACAGCTATTAATGAAGCATATGCTCT	333
Db	241	AATGACACCAAGAGTGTCTTGGGAAAGCAATTAACACAGCTATTAATGAAGCATATGCTCT	300
Qy	334	AAGAGTAAACAAACACAGCTATTAATTAATCAATGATGGAGTCCCTAAACCTGTGAATTAC	393
Db	301	AAGAGTAAACAAACACAGCTATTAATTAATCAATGATGGAGTCCCTAAACCTGTGAATTAC	360
Qy	394	AGTAGAAAGTATGCAACACCAAGGTTTACTGCTGTGGAACAGAGTTCAAGGTTCTGG	453
Db	361	AGTAGAAAGTATGCAACACCAAGGTTTACTGCTGTGGAACAGAGTTCAAGGTTCTGG	420
Qy	454	CTGATTCATTCCTCCTCCTCAGTTTCTCCAAATTCGGAAGCAAGCTATGATTTCCACCC	513
Db	421	CTGATTCATTCCTCCTCCTCAGTTTCTCCAAATTCGGAAGCAAGCTATGATTTCCACCC	480
Qy	514	ACAGGAGACGAAATGGAACAAAGTGGCATCTGCATTAACCTTTCAAGTACCAACAGTATGAG	573
Db	481	ACAGGAGACGAAATGGAACAAAGTGGCATCTGCATTAACCTTTCAAGTACCAACAGTATGAG	540
Qy	574	GCAATAGATTCTCAGTCTTGTGCAACCCCAACGCTCTATAGTCTCTCATCCAGCC	633
Db	541	GCAATAGATTCTCAGTCTTGTGCTGCAACCCCAACGCTCTATAGTCTCTCATCCAGCC	600
Qy	634	ACCTTTACAGAGGCTCATTACATGCCAGCTGTCACAGGCGCAGCTCATCAGAG	693
Db	601	ACCTTTACAGAGGCTCATTACATGCCAGCTGTCACAGGCGCAGCTCATCAGAG	660
Qy	694	ATTCTGGCAGGCTCTCCACACACTTCAGTCGGCCAGGGAACAAAATTCCTCCATTTT	753
Db	661	ATTCTGGCAGGCTCTCCACACACTTCAGTCGGCCAGGGAACAAAATTCCTCCATTTT	720
Qy	754	GCAAGTCGGATTCCTTTCTTGACGACATCTTTGACGCTGGATGCTCAACGGCTGAAG	813
Db	721	GCAAGTCGGATTCCTTTCTTGACGACATCTTTGACGCTGGATGCTCAACGGCTGAAG	780
Qy	814	ACACACTTGTAAACAGAACTGGCAGCGAAAGAACAGAGCTTCCTTCAAACTGCTCC	873
Db	781	ACACACTTGTAAACAGAACTGGCAGCGAAAGAACAGAGCTTCCTTCAAACTGCTCC	840
Qy	874	CTTCTTTACCATGCTACATAATAAAGCAATTAATTAACGACACATCTTTATTTCAAGT	933
Db	841	CTTCTTTACCATGCTACATAATAAAGCAATTAATTAACGACACATCTTTATTTCAAGT	900
Qy	934	TCCTTACAGATCAGCGCAAGTGGTGTATTTCCCAAAAGGCGACCAAAATTCGCTGACA	993
Db	901	TCCTTACAGATCAGCGCAAGTGGTGTATTTCCCAAAAGGCGACCAAAATTCGCTGACA	960
Qy	994	TGATTTGGAGACCTTAATCGGAGTCCACACCAAGCTTTCAGAAAGTGGAGGATTCATTTGT	1053
Db	961	TGATTTGGAGACCTTAATCGGAGTCCACACCAAGCTTTCAGAAAGTGGAGGATTCATTTGT	1020
Qy	1054	ACCCAGAAATGGCAAAATTTACCAAGCAATTTCAAGGATTTAGTATTAATAGAAAGCTGT	1113
Db	1021	ACCCAGAAATGGCAAAATTTACCAAGCAATTTCAAGGATTTAGTATTAATAGAAAGCTGT	1080
Qy	1114	AAGTAA 1119	

```

Db      1081 AAGTAA 1086

RESULT 5
AF333389
LOCUS      Homo sapiens lung-specific DLAD mRNA, complete cds. PRI 22-NOV-2001
DEFINITION
ACCESSION AF333389
VERSION    AF333389.1 GI:17046270
KEYWORDS
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 1029)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            2 (bases 1 to 1029)
            Shiohawa,D. and Tanuma,S.I.
            Isolation and characterization of the DLAD/DLAD genes, which lie
            head-to-head with the genes for urate oxidase
            Biochem. Biophys. Res. Commun. 288 (5), 1119-1128 (2001)
JOURNAL   2155624
MEDLINE   11700027
PUBMED    11700027
REFERENCE  2 (bases 1 to 1029)
            Shiohawa,D. and Tanuma,S.-I.
            Direct Submission
            Submitted (04-JAN-2001) Biochemistry, Science Univ. of Tokyo,
            Ichigaya, Shinjuku-Ku, Tokyo 162-0826, Japan
JOURNAL
FEATURES
source
1. .1029
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
408..869
   /note="expressed only in lung; does not include first two
   exons"
   /codon_start=1
   /product="lung-specific DLAD"
   /protein_id="AAL34448.1"
   /db_xref="GI:17046271"
   /translation="MPQLCTRASSEIPQLRLTLQSAQOKFLHPKSDSFLDDIFA
AWMAORLTHLITETWORKROELPNCSLPYHVYNIKALKLSEHSYFSYQDHAKWCI
SQKGTKNRWTCIGDLNRSHPQAFRGGFTCTQNWQIQYQFQGLVLYIESCK"

CDS
Query Match 72.6%; Score 920.6; DB 9; Length 1029;
Best Local Similarity 99.6%; Pred. No. 1.3e-246;
Matches 923; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      335 AGAGTAACAACACACAGCCTATCTAATATACATGATGGAGTCCCTAAACCTGTGAATTACA 394
      |||
      85 AGAGTAACAACACAGCCTATCTAATATACATGATGGAGTCCCTAAACCTGTGAATTACA 144
      |||
      395 GTAGAAAGTATGGACACACCAAGGTTTACTGCTGTGGAACAGAGTTCAGAGGTTCTGGC 454
      |||
      145 GCAGAAAGTATGGACACACCAAGGTTTACTGCTGTGGAACAGAGTTCAGAGGTTCTGGC 204
      |||
      455 TGATTCATTCATCCCTCAGTTTCTCCAAATTCGGAAGAAGGCTATGATATCCACCCA 514
      |||
      205 TGATTCATTCATCCCTCAGTTTCTCCAAATTCGGAAGAAGGCTATGATATCCACCCA 264
      |||
      515 CAGGAGACGAAATGACAAAGTGGCATCTGCATTAACCTTCAAGTACCAACAGTATGAGG 574
      |||
      265 CAGGAGACGAAATGACAAAGTGGCATCTGCATTAACCTTCAAGTACCAACAGTATGAGG 324
      |||
      575 CAATAGATTCTCAGCTCTTGGTCTGCAACCCCAAGCTATAGTGTCTCCATCCACGCA 634
      |||
      325 CAATAGATTCTCAGCTCTTGGTCTGCAACCCCAAGCTATAGTGTCTCCATCCACGCA 384
      |||
      635 CCTTCCACAGAGCTCAATCAGATGCCCGCCAGCTGTGCACAGGCGCCAGCTCATCAGAGA 694
      |||
      385 CCTTCCACAGAGCTCAATCAGATGCCCGCCAGCTGTGCACAGGCGCCAGCTCATCAGAGA 444
      |||
      695 TTCTCGCAGGCTCTCTACACACATTCAGTGGGCCCGGACAAAAATTCCTCCATTTTG 754
      |||

```

```

Db      445 TTCTGGCAGGCTCCTCACCACACTTCAGTGGGCCCGGACAAAAATTCCTCCATTTTG 504
      |||
      755 CAAAGTCGGATTCCTTTCTTGACGACATCTTTTGACGCTGATGGCTCAACGGCTGAAGA 814
      |||
      505 CAAAGTCGGATTCCTTTCTTGACGACATCTTTTGACGCTGATGGCTCAACGGCTGAAGA 564
      |||
      815 CACACTTGTAAACAGAAACCTGGCAGCGGAAAGACAGAGCTTCCTTCAAACTGCTCCC 874
      |||
      565 CACACTTGTAAACAGAAACCTGGCAGCGGAAAGACAGAGCTTCCTTCAAACTGCTCCC 624
      |||
      875 TTCTTACCATGCTCACAATATAAAAGCAATTAATATCAAGACACTCTTATTTCAAGTT 934
      |||
      625 TTCTTACCATGCTCACAATATAAAAGCAATTAATATCAAGACACTCTTATTTCAAGTT 684
      |||
      935 CTTATCAAGATCAGCGCAAGTGGTGTATTTCCCAAAAGGCGCACAAAAATTCGGTGACAT 994
      |||
      685 CTTATCAAGATCAGCGCAAGTGGTGTATTTCCCAAAAGGCGCACAAAAATTCGGTGACAT 744
      |||
      995 GTATTGGAGACCTAAATCGAGTCCACCAAGGCTTCAGAACTGAGGATTCATTTGTA 1054
      |||
      745 GTATTGGAGACCTAAATCGAGTCCACCAAGGCTTCAGAACTGAGGATTCATTTGTA 804
      |||
      1055 CCCAGAAATGGCAAAATTTACCAAGCATTTCAAGATTTAGTATTATATATGAAAGCTGTA 1114
      |||
      805 CCCAGAAATGGCAAAATTTACCAAGCATTTCAAGATTTAGTATTATATGAAAGCTGTA 864
      |||
      1115 AGTAAACTTGGTGAAAGGACACAGGTACTATCATTTGAAACCTTGACAATGGCTCTTCTT 1174
      |||
      865 AGTAAACTTGGTGAAAGGACACAGGTACTATCATTTGAAACCTTGACAATGGCTCTTCTT 924
      |||
      1175 CCATTACACTTCTTTATATTTTAAAGGCTGTGAATATACCTTAACTGCAATATCACA 1234
      |||
      925 CCATTACACTTCTTTATATTTTAAAGGCTGTGAATATACCTTAACTGCAATATCACA 984
      |||
      1235 AAATAAAACATATTTCTCTCATGTTTA 1261
      |||
      985 AAATAAAACATTTTCTCTCAIGTTAA 1011
      |||

RESULT 6
AF178974
LOCUS      1498 bp mRNA linear ROD 29-NOV-1999
DEFINITION Rattus norvegicus deoxyribonuclease DLAD mRNA, complete cds.
ACCESSION AF178974
VERSION    AF178974.1 GI:6470130
KEYWORDS
SOURCE     Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 1498)
            Tanuma,S. and Shiohawa,D.
            Cloning of a cDNA encoding a rat DNase II-like acid DNase
            Biochem. Biophys. Res. Commun. 265 (2), 395-399 (1999)
JOURNAL   20025354
MEDLINE   10558878
PUBMED    10558878
REFERENCE  2 (bases 1 to 1498)
            Shiohawa,D. and Tanuma,S.
            Direct Submission
            Submitted (19-AUG-1999) Biochemistry, Science University of Tokyo,
            Shinjuku-Ku Ichigaya Funagawara-machi 12, Tokyo 162-0826, Japan
JOURNAL
FEATURES
source
1. .1498
   /organism="Rattus norvegicus"
   /mol_type="mRNA"
   /strain="Wistar"
   /db_xref="taxon:10116"
159..1229
   /codon_start=1
   /product="deoxyribonuclease DLAD"
   /protein_id="AAFI3596.1"
   /db_xref="GI:6470131"

```

/translation="MTAQLKALPILFVALSGVLGTPVISCINEDKRAVDWFAFKL PERTSRGCTGMLDYLIDSTWRTWKSHELLINSRSSLSRLRLEQLYEHNAKNDTAY LIYNDAPASVNSYNGYHAKGLVWNRVQGFLLHSPKPPFPEKGEYFESSRQY AOSGLCTIKVSYNTIDSOLFQFNPIYSCFIPNI FRWELIHPOMCAKSSAKIPS RELTVLOSAQGLNFELHFAKSTFTVDDIIPAAWLAQKLKVLHLLVESWQRKHLPSCSL PHYVNIKAIRGLOPDSHPSHEDHSCWVSTQSRWTCIGDLNRSPhQALRSQGF CSKNRIYIQSFDRLVSHVASCN"		ORIGIN	
Query Match 53.8%; Score 681.8; DB 10; Length 1498; Best Local Similarity 74.1%; Pred. No. 1.1e-179; Matches 891; Conservative 0; Mismatches 307; Indels 5; Gaps 2;			
QY	37	AAACAGAAATGATGCGAAGCTGCTTAAGAACTCCCTTTGCTTTGCTTTCTTTGCGCTC	96
Db	150	AGACAGAAATGACAGCCAGCCTCTAAAGCAGCTCTTCTTTGCTTTCTTTGCTTTG	209
QY	97	TTTGGGGTCTGGGGCAGCAAACTTTCATGAGAAATGAAGAAAGGAAAGCTGTGGAC	156
Db	210	TCCGGAGTCTGGGGACACAGTAATCTCTGTCAATTAACGAGATGGTAAGCCGTGGAC	269
QY	157	TGCTTTACTTTTTTAAAGTTTACCTAAAGACAAACAAAGAAAGTGAGAGACTGGGTTA	216
Db	270	TGCTTTGCTTTTATAAGTTTACCAGAGGACCCAGCAGAGAGGTAACGGGATGGGCTG	329
QY	217	GAGTACTGTACTAGACTCTCACTAGAGAGCTGGAGGAGAGTGAGCACTAATGAAT	276
Db	330	GATTACTGTACTGTACTCTCACTAGAGAGCTGGAGGAGAGTGAGCACTAATGAAT	389
QY	277	GACACCAAGAGTGTTTGGGAAGGACATTAACAGCTATATGAAGCATATGCTCTAAG	336
Db	390	AGCAGCAGGAGCTCCCTGGGAAGGACCTGGAGCAGCTGTATGAGCACAATGCCAAG	449
QY	337	AGTAACAACACAGCTATCTAATATACAAATGAGAGTCCCTAAACCTGTGAATACAGT	396
Db	450	---AATGACACAGCTATCTGATATACAAATGAGAGTCCCTAAACCTGTGAATACAGC	506
QY	397	AGAAAGTATGACACACAAAGGTTTACTGTGTGAGCAGAGTCTCAAGGTTCTGGCTG	456
Db	507	GGAAATTTACGGGATGTCACAAAGGCTCTGTGTATGGAACAGAGTGGGGGTTCTGGCTG	566
QY	457	ATTCAATTCCTCCTCCTAGTTTCTCCCAATTCGGAAGAGGCTATGATATCCACCCACA	516
Db	567	ATTCAATTCCTCCTCCTAGTTTCTCCCAATTCGGAAGAGGCTATGATATCCACCCAGCTG	626
QY	517	GGGAGAGAAATGACAAAGTGGCATCTGCATATCTTCAAGTACCAAGCATATGAGGCA	576
Db	627	GGGAGGCAATATGACAAAGTGGCTCTGTCACTCACTCTCAAAATACAGCCAGTATGAGCA	686
QY	577	ATAGATTTCTAGCTCTTGGTCTGCAACCCCAAGCTCTATAGCTCTCCATCCAGCCACC	636
Db	687	ATAGATTTCTAGCTCTTGGTCTTCCAGCAACATCTACAGCTGTTTCTATCCCAACATC	746
QY	637	TTTCCACGAGGAGCTCATTCAATGCTCCAGCTGTGCAACAGGCGGCTCATCAGAGATT	696
Db	747	TTTCTGCTGGGAATCATCCATGCTCAGATGTGTGCAAGTCCAGTCCCTCAAAGATC	806
QY	697	CCTGGCAGGCTCCTCACACACTTCAGTCTGGGCCAGGACAAATTCCTTCAATTTTGA	756
Db	807	CCTAGCGGGCTCCTCAGTACTTCAGTCAAGCCAGGGTCTAACTCTCTCCATTTTGG	866
QY	757	AAGTCGGATCTTTTCTTGAAGATCTTTCAGCTCTGATGCTCAAGGCTGAAGACA	816
Db	867	AAATCAACTTTTTTATCTAGTATGACATCTTTGAGGCTGGATAGCTCAAGAGTGAAGTG	926
QY	817	CACITGTTTAAAGAAACCTGGCAGGAGAAAGACAAAGAGCTTCTTCAAACTGCTCCCTT	876
Db	927	CAITTTGCTGGTAGAATCTCTGGCAGGAGAAAGAACCCAGAGCTTCTTCAAACTGCTCCCTG	986
QY	877	CCTTACCATGTCTCAATATAAAGCAATTAATATCACGACACTCTTATTTTCAGTTCT	936
Db	987	CCTTACCATGTCTCAACATCAAGGCCAATTCGGGGACCTCTGCGAGTCCGACTTCCCTTCT	1046
QY	937	TATCAAGATCACGCCAAGTGGTGTATTTTCCCAAAAGGGCACAAAATCGCTGGACATGT	996
Db	1047	CATCAGCACCAATTCAAAATGGTGTGTGTTTCCCAAAAAGGACTCTCAGGCTCGTGGACCTGT	1106
QY	997	ATTGAGAGCTTAATTCGGAGTCCACACCAAGCTTTTCAAGAGTGGAGGATTCATTTGTACC	1056
Db	1107	ATTGAGAGCTTAATTCGGAGTCCACATCAAGCTTTTCAAGAGTGGAGGATTCATTTGTTCG	1166
QY	1057	CAGAATTGGCAAAATTTACCAAGCATTTTCAAGGATAGTATTATATATGAAAGCTGTAAG	1116
Db	1167	AGAATTCGGTACATTTACCAAGTATTTGATAGTGTAGTTTCCCATTTATGCTCTCTTAAC	1226
QY	1117	TAACTTTGTTGAAAGACACAGTACTATCATTTGAAACCTTGACAAATCGGTCTTCTTCC	1176
Db	1227	TAACTTTGTTGAAAGTCCACCCACTGCCCTTAAAGACACTGGCTCTCGGACATTTGC	1286
QY	1177	ATTACACC--TTCTTTATATTTTAAAGGCTGTGAATATATATATAACCTGCATATACAA	1234
Db	1287	CTTGGATCTATTCTCCATATGTTTAAAGGTTTCTGAGTGTCTTAACATCTCATCAATAAA	1346
QY	1235	AAA 1237	
Db	1347	AAA 1349	
RESULT 7			
AR432046 1652 bp DNA linear PAT 18-DEC-2003			
Sequence 2 from patent US 6653118.			
AR432046			
VERSION AR432046.1 GI:40194249			
KEYWORDS Unknown.			
SOURCE Unknown.			
ORGANISM Unclassified.			
REFERENCE 1 (bases 1 to 1652)			
AUTHORS Tanuma,S.-i. and Shiohawa,D.			
TITLE Deoxyribonuclease, gene encoding same and use thereof			
JOURNAL Patent: US 6653118-A 2 25-NOV-2003;			
FEATURES Location/Qualifiers			
source 1..1652			
/organism="unknown"			
/mol_type="genomic DNA"			
ORIGIN			
Query Match 53.1%; Score 673.8; DB 6; Length 1652; Best Local Similarity 75.1%; Pred. No. 2e-177; Matches 884; Conservative 0; Mismatches 282; Indels 11; Gaps 3;			
QY	37	AAACAGAAATGATGGGAGAGCTGCTTAAGAACTCTTTGCTTTGCTTTCTTTGCGCTC	96
Db	204	AGATAGAAATGACAGCAAGGCTCTAAGAACAGTCTCTTTCTTCTTTCTTTGCGCTC	263
QY	97	TTTGGGCTGCTGGGGCAGCACTATTTTCATGCAAGATGAAGAGGAGGAGGCTGTGGAC	156
Db	284	TCGCGGCTCTGGGGACACCAAGAACTCTATGCAAGATGAATATGTTGTTGAGCTGTGAC	323
QY	157	TGTTTACTTTTATAGTTTACTCTAAAGACAAACAAAGGAAAGTGGAGAGACTGGGTTA	216
Db	324	TGTTTACTTTTATAGTTTACTCTAAAGGAGTCTAGCAAGGAGTGAAGAGGCGGGCTG	383
QY	217	GAGTACCTGTACCTAGACTCTTACAACTAGAGCTGGAGGAGAGTGAAGCACTAATGAT	276
Db	384	CAGTACCTGTACCTAGACTCTTACCAAGGAGTCTAGCAAGGAGTGAAGAGGCGGGCTG	443
QY	277	GACACCAAGAGTGTTTGGAAGGACATTAACACAGCTATATGAAGCATATGCTCTAAG	336
Db	444	AGCAGCAGAGTGTCTCTGGGGAGGACCTTACAGCATCTGTATGACACATAATCCACG	503
QY	337	AGTAACAACACAGCCTATCTAATATACAAATGATGAGTGGTCCCTTAAACCTGTGAATACAGT	396
Db	504	AATGACA---CAGCCTATCTAATATACACGATGGTGTCTCTGGATCTGTGAATACAGC	560

```
QY 397 AGAAGTATGGAACACCAAGGTTTACTGTGTGGAACAGAGTTCAAGGGTCTTGCTG 456
Db 561 AGACAGTATGGAACATGCAAGGTTCTGTGTATGGAACAGAGGTTCTTGCTG 620
QY 457 ATTCAATCCATCCTCAGTTCCTCCAAATTCGGAAGAGGCTATGATATCCACCCACA 516
Db 621 ATACATCTCTTCCAGTTCCTCCAGTTTC-----ATGGCTATGATGATACCAACCTCG 674
QY 517 GGGAGACGAATGGAACAAAGTGGCATCTGCATTAACATTTCAAGTACACACAGTATGGCA 576
Db 675 GGGAGGCGATGGAACAAACCGCACTGATCACTTCGGATACACCCAGTTTGAGAA 734
QY 577 ATAGATCTCAGCTCTTGTCTGCAACCCCAACGCTATATAGCTCTCCATCCAGCCACC 636
Db 735 ATAGATTTTCAGCTCTTGTCTTACACCAACCAACATCTACAGCTGCTTCATTCOAAGCACC 794
QY 637 TTTCACAGGAGTCTCAATTCACATGCCCCAGCTGTGACACAGGCGCAGCTCATCAGAGATT 696
Db 795 TTTCATCGAACTTATCTACATGCCCCGAGTGTGCCCCAATCTCCAGTTCCCTTAAAGATC 854
QY 697 CTGCGAGGCTCTCTACCAACATCTCAGTGGCCCGAGGACCAAAATTTCTTCCATTTTGA 756
Db 855 CTGTGCGGTACTCTGACTGCACTCAGCCAGGCTTAACTTCGTCCATTTTGA 914
QY 757 AAGTCGGATCTTTCTTGAGACATCTTTGACGCTGAGTGTGAGGCTCAACGGCTGAAAGACA 816
Db 915 AAATCAAGTTTTTATCTGATGACATCTTTACAGGATGATAGCTCAAAAGTTGAAAGACA 974
QY 817 CACTTGTGTAACAGAACCTGGCAGGCAAAAGACAGAGCTTCTTCAAACCTGCTCCCTT 1034
Db 975 CATTTGTTAGCAACAACTGGCAGAAAGAAAGAACAGAGCTTCTTCAAACCTGCTCCCTT 1094
QY 877 CTTTACCATGTCTACAATATAAAAGCAATTAATATACGACACTCTTATTTTCAGTTCT 936
Db 1035 CTTTACCATGTCTACAACATCAGTCCATTTGGGTGATCTTCAAGTCTTACTTCACTTCT 1094
QY 937 TATCAAGATCAGCCAGTGTGTATTTCCAAAAGGGCACCAAAATCGCTGGACATGT 996
Db 1095 CGCCAGAGACCATTCAAAATGTTGTGTTTCCATAAAGGGCTCGCAATCGCTGGACCTGC 1154
QY 997 ATTGAGACCTAAATCGGAGTCCACCAAGCTTTCAGAGTGGAGGATTCATTTGTACC 1056
Db 1155 ATTGAGACCTAAATCGAAGCTTACCAAGCTTTCAGAGTGGAGGATTCATCTGTACA 1214
QY 1057 CAGATTCGCAAAATTTACCAAGCTTTCAGAGTGTATATATATGTAAGAGCTGTAAG 1116
Db 1215 AAGATCACTACATTTACAGGCAATTCATAAATATATCTCCGTATGAGGTTCTGTAAG 1274
QY 1117 TAACTTGGTGAAGGACAGAGTACTATCAATGAAACCTTGAAACCTTGACATGGGTCTTCTCC 1176
Db 1275 TAACTTGGTGAAGGACCAACCTCTGCTCTGAAACACCTGGCACTGGAAACATCTGCG 1334
QY 1177 ATTACACG--TTCTTTATATTTTAAAGCCCTGTGAT 1211
Db 1335 CTGGAATCTGTTCTCCATATTTCAAGGCTCTTGAGT 1371

RESULT 8
BD091714 LOCUS 1652 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel deoxyribonuclease, gene encoding thereof and use thereof.
ACCESSION BD091714
VERSION BD091714.1 GI:22637325
KEYWORDS WO 0112793-A/1.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1652)
Tanuma,S. and Shikawa,D.
Novel deoxyribonuclease, gene encoding thereof and use thereof
Patent: WO 0112793-A 1 22-FEB-2001;
SEIICHI TANUMA,DAISUKE SHIKAWA
```

```
COMMENT OS Mus musculus (mouse)
PN WO 0112793-A/1
PD 22-FEB-2001
PF 01-MAY-2000 WO 2000JP002893
PR 17-AUG-1999 JP 99P 230870
PI SEIICHI TANUMA,DAISUKE SHIKAWA
PC C12N9/22,C12N15/12,C12N5/10,C12N7/01,A61K31/195,A61K48/00 CC
FH Key Location/Qualifiers
FT CDS (213)..(1274).

FEATURES
source
1..1652
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

ORIGIN
Query Match 53.1%; Score 673.8; DB 6; Length 1652;
Best Local Similarity 75.1%; Pred.No.2e-177;
Matches 884; Conservative 0; Mismatches 282; Indels 11; Gaps 3;

QY 37 AAACAGAAATGATGGCAAGACTGTGAAGAAATCCTTTGCTTTGCTTCTTCTTGGCCTC 96
Db 204 AGATAGAAATGACAGCAAAAGCTCTAAGACAGTTCTTTTCTTCTTCTTCTTCTTCTT 263
QY 97 TTTGGGTCCTGGGCGACGACAAATTCATGTCAGAAATGAAGAGGAAAGCTGTGGAC 156
Db 264 TCTGGGTCCTGGGCGACGACCAAAATCTCATGCAGAAATGAATATGTTGAAGCTGTGGAC 323
QY 157 TGGTTTACTTTTATAAGTTTACCTTAAAGACAAACAAAGAAAGTGAAGAGAGCTGGTTA 216
Db 324 TGGTTTACTTTTATAAGTTTACCTTAAAGAGGATAGCAAGAGTGAAGAGGCGGCTG 383
QY 217 GAGTACCTGTACTAGACTCTACAACTAGAACTGGAGAGAGTGAAGCACTAATGAAT 276
Db 384 CAGTACCTGTACTAGACTCTACAACTAGAACTGGAGAGAGCTCTACCTGATTAAC 443
QY 277 GACACCAAGAGTGTGTTGGGAAGGACATTACAAAGCTATATGAAGCATATCCTCTAAG 336
Db 444 AGCACCAAGAGTGTGTTGGGAAGGACCTTACAGCATCTGTATGACACATATTTCCACG 503
QY 337 AGTAAACAACAGAGCTTACTTAAATATACATGATGAGTCCCTTAAACCTGTGATTAAGT 396
Db 504 AATGACA--CAGCTTATCTAATATACAAAGTGTGCTTGGATCTGTGATTAAGT 560
QY 397 AGAAAGTATGACACACAAAGGTTTACTGTGTGAAACAGAGTTTCAAGGGTCTGCGCTG 456
Db 561 AGACAGTATGACATGCAAGGCTGCTGTGTATGGAACAGACGCGAGGTTCTGCGCTG 620
QY 457 ATTCAATTCATCCCTCAGTTTCTTCCAAATTCGGAAGAGGCTATGATTTCCACCCACA 516
Db 621 ATACACTCTGTTCCTCAAGTTTCCCTCAGTTT-----ATGGCTATGATGATCCCACTCG 674
QY 517 GGGAGACGAATGGAACAAAGTGGCATCTGCATTAACATTTCAAGTACACACAGTATGAGCA 576
Db 675 GGGAGGCGATGGAACAAACCGCACTGATCACTTCGGATACACCCAGTTTGAGAA 734
QY 577 ATAGATCTCAGCTCTTGTGTCACACCTTCTATAGCTCTTCTTCTTCTTCTTCTTCTTCT 636
Db 735 ATAGATTTTCAGCTCTTGTGTTTACACCAACCAACATCTACAGCTGCTTCTTCTTCTTCT 794
QY 637 TTTCACAGAGGCTCATTCATGCCCCAGCTGTGCAACCGGCGCAGCTCATCAGAGATT 696
Db 795 TTTCACAGAGGCTCATTCATGCCCCAGCTGTGCAACCTGCTTCTTCTTCTTCTTCTTCTT 854
QY 697 CTGCGAGGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 756
Db 855 CTGTCGCGTACCTCTGCTGAACTGCACTCAGCCAGGCTTAACTTCTGCTCCATTTTGA 914
QY 757 AAGTCGGATCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 816
Db 915 AAATCAAGTTTTTATCTGATGACATCTTTACAGGATGATGATGATGATGATGATGATGAT 974
QY 817 CACTTGTGTAACAGAACCTGGCAGGCAAAAGACAGAGCTTCTTCAAACCTGCTCCCTT 876
```

RESULT	9
AFI28888	
LOCUS	AFI28888 1654 bp mRNA linear ROD 06-DEC-1999
DEFINITION	Mus musculus deoxyribonuclease DLAD mRNA, complete cds.
ACCESSION	AFI28888
VERSION	AFI28888.1 GI:6175549
KEYWORDS	
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1654) Shiokawa,D. and Tanuma,S. DLAD,a novel mammalian divalent cation-independent endonuclease with homology to DNase II Nucleic Acids Res. 27 (20), 4083-4089 (1999)
REFERENCE	99428663 10497274
PUBMED	
REFERENCE	2 (bases 1 to 1654) Shiokawa,D. and Tanuma,S. Direct Submission
AUTHORS	
TITLE	Submitted (16-FEB-1999) Biochemistry, Science Univ. of Tokyo, Shinjuku-ku Ichigaya Funagawara-machi 12, Tokyo 162-0826, Japan
JOURNAL	Location/Qualifiers
JOURNALS	1. .1654 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6" /db_xref="taxon:10090" /tissue_type="liver" 213..1277 /codon_start=1 /product="deoxyribonuclease DLAD" /protein_id="AAF05082.1" /db_xref="GI:6175550" /translation="MTAPRLRTVLSLFPALSGVLGTPEISCRNEYGRAVDWFIYKIK PKRTSKEEAGQLYLIDSTRQTNWKSLYILNISRSLGRTLQHLYPTNSTNDTA LIYNDGPGSVSNQSYGHAKGLTWNRSTQGFWLIHSPKFFPHGYEYPPISGRRYA TGICITGFQSFDEEDFOLLVLPNIYSCTPFSFTFKWLKIYMPRCANSSSKLIPVR LAELHSACLNLHFPAKSIFYDDIFTGMIAQKLTHLLAQTWKKQDELPSNCSLP HYVNKTSGTVSKSYFSERODSHKWCVSIKGSANEWTCTIGDILNRSIHQAIRGGGFCIC KNHIYQAFHKLYLRYGCK"
FEATURES	
source	
CDS	
ORIGIN	53.1%; Score 673.8; DB 10; Length 1654;
Query Match	





mrna join(<1567..1691,4903..5080,11439..11520,13835..13996,15315..15512,17497..>.17837)  
/gene="DLAD"  
/product="endonuclease DLAD"  
CDS join(1567..1691,4903..5080,11439..11520,13835..13996,15315..15512,17497..17837)  
/gene="DLAD"  
/codon\_start=1  
/product="endonuclease DLAD"  
/protein\_id="AAL34449.1"  
/db\_xref="GI:17046273"  
translation="MGKQWRLRTSFALLFLGLFVLAATISCRNEEGKAVDMFT  
FYKLPRQKESGETGLVLYLDSTTRSWRKESOLMNDTKSVLGRLOQLYEAAYS  
NNFAYLYNDGVKPNVSRKYKHTGLLLNVEVQFWLIHSIPQPPPIPEEGYDVP  
TGRNQSOGCITFRKQVFAIDSQLVNDPNVYSCSIPTAFHOELHNPQLCTRASS  
SETPGRLLTLQSAQOKFLHFAKSDFLDDIFAAWMAQRKLHLLTETWQRQELP  
SNGSLPHYVNIKAIKLSRHSYFSSYQDHAKWCISQKTKNRWTCIGDLNRSHPQAFR  
SGGFICTQNWQIYQAFQGLVLYYESCK"  
  
ORIGIN  
Query Match 38.5%; Score 487.8; DB 9; Length 19164;  
Best Local Similarity 99.6%; Pred. No. 3.2e-125;  
Matches 489; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 778 GACATCTTTTCAGCTGGATGGCTCAACGGCTGAAGACACACTTTGTTAAGAGAACTGG 837  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
17496 GACATCTTTTCAGCTGGATGGCTCAACGGCTGAAGACACACTTTGTTAAGAGAACTGG 17555  
  
QY 838 CAGCGAAAAGACAGAGCTTCCCTTCAAACTGCTCCCTTCTTCAACGATGCTCAATATA 897  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
17556 CAGCGAAAAGACAGAGCTTCCCTTCAAACTGCTCCCTTCTTCAACGATGCTCAATATA 17615  
  
QY 898 AAAGCAATTAAATATCAAGCACTCTTATTTCACTTCTTATCAAGATCAGCACTGG 957  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
17616 AAAGCAATTAAATATCAAGCACTCTTATTTCACTTCTTATCAAGATCAGCACTGG 17675  
  
QY 958 TGATTTTCCCAAGAGGACCAAAATCGCTGGACATGATTTGAGACCTTAAATCGGAGT 1017  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
17676 TGATTTTCCCAAGAGGACCAAAATCGCTGGACATGATTTGAGACCTTAAATCGGAGT 17735  
  
QY 1018 CCACACCAAGCCTTCAGAGTGGAGGATTCATTTGTACCCAGAAATGGCAATTTACCA 1077  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
17736 CCACACCAAGCCTTCAGAGTGGAGGATTCATTTGTACCCAGAAATGGCAATTTACCA 17795  
  
QY 1078 GCATTTCAGATTAGTATTTATCTATCAAGCTGTAGTAACTTGTGTGAAGAGACACA 1137  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
17796 GCATTTCAGATTAGTATTTATCTATCAAGCTGTAGTAACTTGTGTGAAGAGACACA 17855  
  
QY 1138 GGTACTATCATTTGAAACCTTGACATGGGTCTTCTTCCATTTACACCTTCTTTATATTT 1197  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
17856 GGTACTATCATTTGAAACCTTGACATGGGTCTTCTTCCATTTACACCTTCTTTATATTT 17915  
  
QY 1198 AAAGGCGCTGTGAATATCTTTAATACCTGCATATCAAAAATAAATTTCTCTCATG 1257  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
17916 AAAGGCGCTGTGAATATCTTTAATACCTGCATATCAAAAATAAATTTCTCTCATG 17975  
  
QY 1258 TTTACCATTTA 1268  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
17976 TTTACCATTTA 17986

RESULT 12  
AC027606  
LOCUS AC027606 164529 bp DNA linear HTG 15-JUL-2000  
DEFINITION Homo sapiens chromosome 1 clone RP11-30M11, WORKING DRAFT SEQUENCE,  
18 unordered pieces.  
ACCESSION AC027606  
VERSION AC027606.4 GI:9211324  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 164529)  
Waterston,R.H.  
The sequence of Homo sapiens clone  
Unpublished  
2 (bases 1 to 164529)  
Waterston,R.H.  
Direct Submission  
Submitted (30-MAR-2000) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
On Jul 15, 2000 this sequence version replaced gi:8844157.

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc/index.shtml  
----- Project Information -----  
Center project name: H\_NH0030M11  
----- Summary Statistics -----  
Sequencing vector: M13; 100%  
Sequencing vector: plasmid; 0%  
Chemistry: Dye-primer ET; 100% of reads  
Chemistry: Dye-terminator Big Dye; 0% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 156994 bases at least Q40  
Consensus quality: 159185 bases at least Q30  
Consensus quality: 160492 bases at least Q20  
Insert size: 177000; agarose-fp  
Insert size: 162829; sum-of-contigs  
Quality coverage: 4.41 in Q20 bases; agarose-fp  
Quality coverage: 4.23 in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 18 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1304: contig of 1304 bp in length  
1404: gap of unknown length  
2815: contig of 1411 bp in length  
2915: gap of unknown length  
6065: contig of 3150 bp in length  
6165: gap of unknown length  
9367: contig of 3202 bp in length  
9467: gap of unknown length  
13313: contig of 3846 bp in length  
13413: gap of unknown length  
16861: contig of 3448 bp in length  
16961: gap of unknown length  
20927: contig of 3966 bp in length  
21027: gap of unknown length  
25571: contig of 4544 bp in length  
25671: gap of unknown length  
31485: contig of 5814 bp in length  
31585: gap of unknown length  
37141: contig of 5556 bp in length  
37241: gap of unknown length  
44119: contig of 6878 bp in length  
44219: gap of unknown length  
51993: contig of 7774 bp in length  
52093: gap of unknown length  
61678: contig of 9585 bp in length  
61778: gap of unknown length  
73742: contig of 11964 bp in length  
73842: gap of unknown length  
90362: contig of 16520 bp in length  
90462: gap of unknown length  
110115: contig of 19653 bp in length  
110215: gap of unknown length  
129841: contig of 19626 bp in length



```

* 129842 129941: gap of unknown length
* 129942 164529: contig of 34588 bp in length.
FEATURES
    source
        1..164529
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="1"
            /clones="RP11-30M11"
            1..1304
                /note="assembly_name:Contig11"
            1405..2815
                /note="assembly_name:Contig12"
            clone_end:SP6
            vector_side:right
            2916..6065
                /note="assembly_name:Contig13"
            6166..9367
                /note="assembly_name:Contig14"
            9468..13313
                /note="assembly_name:Contig15"
            13414..16861
                /note="assembly_name:Contig16"
            16962..20927
                /note="assembly_name:Contig17"
            21028..25571
                /note="assembly_name:Contig18"
            25672..31485
                /note="assembly_name:Contig19"
            31586..37141
                /note="assembly_name:Contig20"
            37242..44119
                /note="assembly_name:Contig21"
            44220..51993
                /note="assembly_name:Contig22"
            52094..61678
                /note="assembly_name:Contig23"
            61779..73742
                /note="assembly_name:Contig24"
            73843..90362
                /note="assembly_name:Contig25"
            90463..110115
                /note="assembly_name:Contig26"
            110216..129841
                /note="assembly_name:Contig27"
            129942..164529
                /note="assembly_name:Contig28"

ORIGIN
Query Match      38.5%; Score 487.8; DB 2; Length 164529;
Best Local Similarity 99.6%; Pred. No. 3.8e-125;
Matches 489; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 778 GACATCTTTGACGCTGGATGCTCAACGGCTGACGACACATCTGTTAAAGAAACCTGG 837
Db 143110 GACATCTTTGACGCTGGATGCTCAACGGCTGACGACACATCTGTTAAAGAAACCTGG 143169

QY 838 CAGCGAAAAGACAGAGCTTCCTTCAAACTGCTCCCTTCCCTTACCATGCTACAAATATA 897
Db 143170 CAGCGAAAAGACAGAGCTTCCTTCAAACTGCTCCCTTCCCTTACCATGCTACAAATATA 143229

QY 898 AAAGCAATTAATATCATCGACATCTTATTTTACGTTCTTATCAAGATCAACGCAAGTGG 957
Db 143230 AAAGCAATTAATATCATCGACATCTTATTTTACGTTCTTATCAAGATCAACGCAAGTGG 143289

QY 958 TGTATTTCCCAAGGGCACCACCAAAATCGCTGGACATGATTTGGAGACTAAATCGGAGT 1017
Db 143290 TGTATTTCCCAAGGGCACCACCAAAATCGCTGGACATGATTTGGAGACTAAATCGGAGT 143349

QY 1018 CCACCAAGGCTTCAGAAAGTGGAGGATTCATTTGTACCCAGAAATGGCAATTTACCAA 1077
Db 143350 CCACCAAGGCTTCAGAAAGTGGAGGATTCATTTGTACCCAGAAATGGCAATTTACCAA 143409

QY 1078 GCATTTCAAGGATTAGTATTATATCTATGAAAGCTGTAAAGTAAACCTTGTGTAAGACACA 1137
Db 143410 GCATTTCAAGGATTAGTATTATATCTATGAAAGCTGTAAAGTAAACCTTGTGTAAGACACA 143469

QY 1138 GGTACTATCATTTGAAAAACCTTGACAAATGGGTCTTCTCCATTACACCTTCTTTATATTTT 1197
Db 143470 GGTACTATCATTTGAAAAACCTTGACAAATGGGTCTTCTCCATTACACCTTCTTTATATTTT 143529

QY 1198 AAGGCGCTGTGATATATCTTATTAACCTGCAATATACAAAATAAAACATATTTCTCTCATG 1257
Db 143530 AAGGCGCTGTGATATATCTTATTAACCTGCAATATACAAAATAAAACATATTTCTCTCATG 143589

QY 1258 TTTACCAATTTA 1268
Db 143590 TTTACCAATTTA 143600

RESULT 13
AL359273
LOCUS
DEFINITION
    Human DNA sequence from clone RP11-376N17 on chromosome 1, complete
    sequence.
ACCESSION
    AL359273
VERSION
    AL359273.11 GI:12191184
KEYWORDS
    HTG.
SOURCE
    Homo sapiens (human)
    ORGANISM
        Homo sapiens
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
        Wallis,J.
REFERENCE
    1 (bases 1 to 205034)
AUTHORS
    Direct Submission
JOURNAL
    Submitted (10-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
    CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
    requests: clonerequest@sanger.ac.uk
    On Jan 13, 2001 this sequence version replaced gi:12043432.
    During sequence assembly data is compared from overlapping clones.
    Where differences are found these are annotated as variations
    together with a note of the overlapping clone name. Note that the
    variation annotation may not be found in the sequence submission
    corresponding to the overlapping clone, as we submit sequences with
    only a small overlap as described above.
    This sequence has been finished according to sequence map criteria
    as follows. An attempt is made to resolve all sequencing problems,
    such as compressions and repeats, but not necessarily within known
    annotated repeat sequence elements. Where the sequence is
    ambiguous, there is an annotation using the 'unsure' feature key.
    The following abbreviations are used to associate primary accession
    numbers given in the feature table with their source databases:
    Emi, EMBL; Swi, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information
    on the WORMPEP database can be found at
    http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
    was generated from part of bacterial clone contigs of human
    chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
    Group. Further information can be found at
    http://www.sanger.ac.uk/HGP/Chri
    RP11-376N17 is from the library RPCI-11.2 constructed by the group
    of Pieter de Jong. For further details see
    http://www.chori.org/bacpac/home.htm
    VECTOR: pBAC3.6
    This sequence is the entire insert of clone RP11-376N17 The true
    right end of clone RP11-82H13 is at 76245 in this sequence.
FEATURES
    source
        1..205034
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="1"
            /clones="RP11-376N17"
            /clone_lib="RPCI-11.2"
            512..622
                /note="L2 repeat: matches 2588..2704 of consensus"
            764..845
                /note="L2 repeat: matches 2588..2704 of consensus"
                repeat_region
                repeat_region

```

```

/note="41 copies 2 mer ct 68% conserved"
repeat_region 2021. .2169
/note="MIR repeat: matches 58. .210 of consensus"
repeat_region 2802. .2886
/note="MIR repeat: matches 102. .192 of consensus"
repeat_region 3607. .4457
/note="LIMB3 repeat: matches 5284. .6181 of consensus"
repeat_region 4671. .4952
/note="AluSx repeat: matches 1. .301 of consensus"
repeat_region 6766. .7048
/note="AluSx repeat: matches 1. .307 of consensus"
repeat_region 7636. .7693
/note="29 copies 2 mer at 87% conserved"
repeat_region 8235. .8452
/note="MER30 repeat: matches 1. .230 of consensus"
repeat_region 8483. .8592
/note="MIR repeat: matches 67. .166 of consensus"
repeat_region 8578. .8717
/note="MIR repeat: matches 118. .257 of consensus"
repeat_region 8897. .9474
/note="L2 repeat: matches 1728. .2356 of consensus"
repeat_region 9508. .10010
/note="MLTIF repeat: matches 1. .541 of consensus"
repeat_region 10247. .10349
/note="MLTID repeat: matches 23. .120 of consensus"
repeat_region 10351. .10453
/note="LIMB6 repeat: matches 6074. .6176 of consensus"
repeat_region 10457. .10693
/note="MER7A repeat: matches 1. .223 of consensus"
repeat_region 10694. .10817
/note="LIMB7 repeat: matches 6159. .6283 of consensus"
repeat_region 10838. .10883
/note="23 copies 2 mer ta 76% conserved"
repeat_region 10908. .11211
/note="MLTID repeat: matches 192. .505 of consensus"
repeat_region 11215. .11380
/note="L2 repeat: matches 2555. .2749 of consensus"
repeat_region 11742. .11813
/note="MIR repeat: matches 65. .145 of consensus"
repeat_region 11838. .12325
/note="L2 repeat: matches 2245. .2746 of consensus"
repeat_region 12326. .12658
/note="MER33 repeat: matches 5. .324 of consensus"
repeat_region 12659. .12770
/note="L2 repeat: matches 2131. .2245 of consensus"
repeat_region 12771. .13068
/note="AluY repeat: matches 1. .294 of consensus"
repeat_region 13069. .13324
/note="L2 repeat: matches 1895. .2131 of consensus"
repeat_region 13410. .13524
/note="L2 repeat: matches 2269. .2384 of consensus"
repeat_region 13668. .14270
/note="MER67B repeat: matches 1. .661 of consensus"
repeat_region 14405. .14569
/note="MER69A repeat: matches 1. .163 of consensus"
repeat_region 14842. .16800
/note="LIMC1 repeat: matches 4339. .6327 of consensus"
repeat_region 16828. .16920
/note="MSTD repeat: matches 41. .132 of consensus"
repeat_region 17020. .17238
/note="MSTD repeat: matches 199. .426 of consensus"
repeat_region 17239. .17656
/note="LIMC1 repeat: matches 3954. .4343 of consensus"
repeat_region 17650. .18127
/note="LIMC4 repeat: matches 329. .831 of consensus"
repeat_region 18128. .18405
/note="AluDb repeat: matches 3. .278 of consensus"
repeat_region 18406. .18515
/note="LIMC4 repeat: matches 224. .329 of consensus"
repeat_region 18516. .18600
/note="MER5A repeat: matches 67. .162 of consensus"
repeat_region 20386. .20677
/note="AluSg repeat: matches 1. .300 of consensus"

```

```

21518. .21560
/note="L2 repeat: matches 2624. .2667 of consensus"
repeat_region 21587. .21746
/note="MIR repeat: matches 67. .231 of consensus"
repeat_region 21847. .22101
/note="L2 repeat: matches 2244. .2489 of consensus"
repeat_region 22265. .24015
/note="LIMC1 repeat: matches 2197. .3954 of consensus"
repeat_region 24177. .24480
/note="AluSg repeat: matches 6. .301 of consensus"
repeat_region 24979. .25248
/note="MLTII repeat: matches 63. .321 of consensus"
repeat_region 25329. .25560
/note="HALI repeat: matches 557. .792 of consensus"
repeat_region 25590. .26037
/note="LIMEC repeat: matches 1207. .1648 of consensus"
repeat_region 27307. .27752
/note="MLTIC repeat: matches 1. .466 of consensus"
repeat_region 28933. .29100
/note="HALI repeat: matches 585. .772 of consensus"
repeat_region 29101. .30182
/note="MER11C repeat: matches 1. .1071 of consensus"
repeat_region 30183. .30284
/note="HALI repeat: matches 772. .868 of consensus"
repeat_region 30670. .31206
/note="MLTIF repeat: matches 1. .538 of consensus"
repeat_region 31207. .31928
/note="MER67B repeat: matches 2. .661 of consensus"
repeat_region 31931. .32036
/note="53 copies 2 mer aa 63% conserved"
repeat_region 32476. .32893
/note="MSTA repeat: matches 2. .425 of consensus"
repeat_region 33144. .33452
/note="AluSx repeat: matches 1. .310 of consensus"
repeat_region 34297. .34983
/note="L2 repeat: matches 1992. .2750 of consensus"
repeat_region 35174. .35467
/note="AluSg repeat: matches 3. .297 of consensus"
repeat_region 35900. .36248
/note="THEIC repeat: matches 3. .371 of consensus"
repeat_region 36615. .36696
/note="L2 repeat: matches 1711. .1779 of consensus"
repeat_region 36697. .39050
/note="TIGER1 repeat: matches 3. .2418 of consensus"
repeat_region 39051. .39300
/note="L2 repeat: matches 1779. .2094 of consensus"
repeat_region 39347. .39578
/note="MIR repeat: matches 23. .262 of consensus"
repeat_region 39897. .39988
/note="AluSg/x repeat: matches 218. .309 of consensus"
repeat_region 40510. .41307
/note="CpG island"
repeat_region 41742. .41781
/evidence=not_experimental
repeat_region 43592. .44272
/note="match: STS: Em:G51282"
repeat_region 44412. .44473
/note="31 copies 2 mer tt 67% conserved"
repeat_region 44611. .44660
/note="25 copies 2 mer ca 86% conserved"
repeat_region 46419. .46584
/note="MER3 repeat: matches 1. .164 of consensus"
repeat_region 46591. .48600
/note="LIMC1 repeat: matches 4097. .6163 of consensus"
repeat_region 47717. .47748
/note="Sequence from overlapping clone bA82H13 (AL450063).
Assembly confirmed by restriction digest data."
repeat_region 48601. .48916
/note="AluSg repeat: matches 1. .287 of consensus"
repeat_region 48922. .48967

```

Query Match 38.5%; Score 487.8; DB 9; Length 205034;  
 Best Local Similarity 99.6%; Pred. No. 3.8e-125;  
 Matches 489; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 778 GACATCTTTGACGCTGATGGCTCAACGGCTGAAGACACACTTGTGTTAACGAAACCTGG 837  
 Db 156611 GACATCTTTGACGCTGATGGCTCAACGGCTGAAGACACACTTGTGTTAACGAAACCTGG 156670

QY 838 CAGCGAAAAGACAAAGAGCTTCCCTTCAAACTGCTCCCTTCCCTTACCAATGTTACAATATA 897  
 Db 156671 CAGCGAAAAGACAAAGAGCTTCCCTTCAAACTGCTCCCTTCCCTTACCAATGTTACAATATA 156730

QY 898 AAGCAATTAATATCAAGCACTCTTATTTCAAGTCTTATCAAGATCAAGCAAGTGG 957  
 Db 156731 AAGCAATTAATATCAAGCACTCTTATTTCAAGTCTTATCAAGATCAAGCAAGTGG 156790

QY 958 TGATTTTCCCAAAGGCGACCAAAATCGCTGGACATGTATGAGACCTTAATFCGAGT 1017  
 Db 156791 TGATTTTCCCAAAGGCGACCAAAATCGCTGGACATGTATGAGACCTTAATFCGAGT 156850

QY 1018 CCACACCAAGCCTTCAGAACTGGAGGATTCATTTGTACCCAGAAATGGCAATTTACCAA 1077  
 Db 156851 CCACACCAAGCCTTCAGAACTGGAGGATTCATTTGTACCCAGAAATGGCAATTTACCAA 156910

QY 1078 GCATTTTCAAGGATGATTTATATCTATGAAGCTGTAAGTAACTTGGTGAAAGGACACA 1137  
 Db 156911 GCATTTTCAAGGATGATTTATATCTATGAAGCTGTAAGTAACTTGGTGAAAGGACACA 156970

QY 1138 GGTACTATCATTCGAAACCTTGCAATGGCTTCTTCCATTCACCTTCTTTATATTTT 1197  
 Db 156971 GGTACTATCATTCGAAACCTTGCAATGGCTTCTTCCATTCACCTTCTTTATATTTT 157030

QY 1198 AAGGCTCTGATATATCTATTAACCTGCAATACAAATATAAAATATTTCTCTCATG 1257  
 Db 157031 AAGGCTCTGATATATCTATTAACCTGCAATACAAATATAAAATATTTCTCTCATG 157090

QY 1258 TTTACCATTTA 1268  
 Db 157091 TTTACCATTTA 157101

RESULT 14  
 AF33460386  
 LOCUS  
 DEFINITION Mus musculus deoxyribonuclease DLAD (Dlad) gene, exon 6 and complete cds.  
 ACCESSION AF334608  
 VERSION AF334608.1 GI:17046280  
 KEYWORDS  
 SEGMENT 6 of 6  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 2337)  
 Shikawa, D. and Tanuma, S.I.  
 Isolation and characterization of the DLAD/Dlad genes, which lie head-to-head with the genes for urate oxidase  
 Biochem. Biophys. Res. Commun. 288 (5), 1119-1128 (2001)  
 21556924  
 PUBMED 11700027  
 REFERENCE 2 (bases 1 to 2337)  
 Shikawa, D. and Tanuma, S.-I.  
 Direct Submission  
 TITLE Submitted (09-JAN-2001) Department of Biochemistry, Science University of Tokyo, Ichigaya, Shinjuku-ku, Tokyo 162-0826, Japan  
 FEATURES  
 Location/Qualifiers  
 1..2337  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"

gene  
 order (AF334603.1:1476.1825, AF334604.1:1.1374, AF334605.1:1.1089, AF334606.1:1.719, AF334607.1:1.642, 1..1182)  
 /gene="Dlad"  
 mRNA  
 join (AF334603.1:1476.1588, AF334604.1:665.842, AF334605.1:630.708, AF334606.1:484.639, AF334607.1:102.299, 486.1182)  
 /gene="Dlad"  
 /product="deoxyribonuclease DLAD"  
 CDS  
 join (AF334603.1:1476.1588, AF334604.1:665.842, AF334605.1:630.708, AF334606.1:484.639, AF334607.1:102.299, 486.826)  
 /gene="Dlad"  
 /codon\_start=1  
 /product="deoxyribonuclease DLAD"  
 /protein\_id="AAL34450.1"  
 /db\_xref="GI:17046281"  
 /translation="MTAKPLRTVLSLLFFALSGVLGTPETISCRNEYGEAVDMTFYKL  
 PKTSKASEAGLQYLYLDSTROTWNKSLYLINLSALGRTIQLHYDTHSNDTAY  
 LIYDGVPSVNSYROYHAKGLLVNRTQGFWLIIHSVPKPPVHGVEYTSGRVYQ  
 TGICTFGVSQFEEDFOLLVLPNIYSCFIPSTFWHLTYMPRMCAANSKIPVRY  
 LAELHGAQGLNFVHFAKSFYTDIDFTGMLAKLTHLLAOTWQKKQLPNSCSLPY  
 KNIKISIGVTSKVSFSSRODHSKWCVSIGSANRWTCTGDLNRSLSHQLRGSGFCT  
 KHHYIYQAFHKLILRYGCK"  
 486.1182  
 /gene="Dlad"  
 /number=6  
 exon  
 ORIGIN  
 Query Match 20.0%; Score 254; DB 10; Length 2337;  
 Best Local Similarity 75.5%; Pred. No. 9.4e-60;  
 Matches 329; Conservative 0; Mismatches 105; Indels 2; Gaps 1;

QY 778 GACATCTTTGACGCTGATGGCTCAACGGCTGAAGACACACTTGTGTTAACGAAACCTGG 837  
 Db 485 GACATCTTTACAGATGATGATGCTCAAAAGTTGAAGACACATTTGTAGCACAACCTGG 544

QY 838 CAGCGAAAAGACAAAGAGCTTCCCTTCAAACTGCTCCCTTCCCTTACCAATATA 897  
 Db 545 CAGCGAAAAGACAAAGAGCTTCCCTTCAAACTGCTCCCTTCCCTTACCAATATA 604

QY 898 AAGCAATTAATATCAAGCACTCTTATTTCAAGTCTTATCAAGATCAAGCAAGTGG 957  
 Db 605 AAGTCCATTTGGGGTAACTTCCAAAGTCTTACTTCAAGTCTTCCGCAAGACCATTCCTCAATGG 664

QY 958 TGATTTTCCCAAAGGCGACCAAAATCGCTGGACATGTATGAGACCTTAATFCGAGT 1017  
 Db 665 TGATTTTCCCAAAGGCGCTCCGCAATCGTGGACCTGCAATGAGACCTTAATGGAAGC 724

QY 1018 CCACACCAAGCCTTCAGAACTGGAGGATTCATTTGTACCCAGAAATGGCAATTTACCAA 1077  
 Db 725 CTACACCAAGCCTTAAGAGTGGAGGATTCATCTGTACAAAGAAATCACTACATTTACCAAG 784

QY 1078 GCATTTTCAAGGATGATTTATATCTATGAAGCTGTAAGTAACTTGGTGAAAGGACACA 1137  
 Db 785 GCATTTTCAATTAATATATCTCCGTTATGGGTCTGTGTAAGTAACTTCGCTGTAAGGCGCAC 844

QY 1138 GGTACTATCATTCGAAACCTTGCAATGGCTTCTTCCATTCACCTTCTTTATATTT 1195  
 Db 845 CCTCTGCTGCTTGAACACACTGGCACTGGAACATCTCGCTTGGATCTGTCTTCATTAAT 904

QY 1196 TTAAGGCTCTGTAAT 1211  
 Db 905 TTCAAGGCTCTGAGT 920

RESULT 15  
 AC127571  
 LOCUS  
 DEFINITION Mus musculus chromosome UNK clone RP24-227D4, WORKING DRAFT  
 SEQUENCE, 4 unoriented pieces.  
 ACCESSION AC127571  
 VERSION AC127571.2 GI:24080752

197457 bp DNA linear HTG 17-OCT-2002

HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
Bukaryota; Metazoa; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 197457)  
McPherson, J.D. and Waterston, R.H.  
The sequence of Mus musculus clone  
Unpublished  
2 (bases 1 to 197457)  
McPherson, J.D. and Waterston, R.H.  
Direct Submission  
Submitted (17-JUL-2002) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
3 (bases 1 to 197457)  
McPherson, J.D. and Waterston, R.H.  
Direct Submission  
Submitted (17-OCT-2002) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
On Oct 17, 2002 this sequence version replaced gi:21887002.  
----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc/index.shtml  
Contact: submissions@wustl.edu  
----- Project Information -----  
Center project name: M BB0227D04  
----- Summary Statistics -----  
Sequencing vector: M13; 0%  
Chemistry: Dye-terminator; 100%  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 195668 bases at least Q40  
Consensus quality: 196117 bases at least Q30  
Consensus quality: 196344 bases at least Q20  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 4 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 31865: contig of 31865 bp in length  
\* 31866 31865: gap of unknown length  
\* 31966 70958: contig of 38993 bp in length  
\* 70959 71058: gap of unknown length  
\* 71059 136518: contig of 65460 bp in length  
\* 136519 136618: gap of unknown length  
\* 136619 197457: contig of 60839 bp in length.  
FEATURES  
source  
1..197457  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="UNK"  
/clone="RP24-227D4"  
1..31865  
/note="assembly\_name:Contig14"  
31966..70958  
/note="assembly\_name:Contig15"  
71059..136518  
/note="assembly\_name:Contig16"  
136619..197457  
/note="assembly\_name:Contig17"  
ORIGIN  
Query Match 20.0%; Score 254; DB 2; Length 197457;  
Best Local Similarity 75.5%; Pred. No. 1.3e-59;

Matches 329; Conservative 0; Mismatches 105; Indels 2; Gaps 1;  
QY 778 GACATCTTTGACGCTGGATGGCTCAACGGCTGAGACACACTGTGTAAAGAAACCTGG 837  
Db 68107 GACATCTTTTACAGGATGGATAGCTCAAAGTTGAAGACACATTTGTTAGCACAACCTGG 68166  
QY 838 CAGCGAAAAAGACAGAGCTTCTTCAAACCTGCTCCCTTCCCTTACCATGCTCTACATATA 897  
Db 68167 CAGAAAAAGAAACAGAGCTTCTTCAAACCTGCTCCCTTACCATGCTCTACATATA 68226  
QY 898 AAGCAATTAATATACAGACACTCTTATTCAGTTCTTATCAAGATCACGCCAAGTGG 957  
Db 68227 AAGTCCATTGGGGTAACTTCCAAAGTCTTACTTCAGTTCTCGCCCAAGACCAATTC 68286  
QY 958 TGTATTTTCCCAAAAGGCGACCAAAATCGCTGACATGTATTTGGAGACCTTAAATCGGAGT 1017  
Db 68287 TGTATTTTCCCAAAAGGCGCTCCGCAATCGCTGACATGTATTTGGAGACCTTAAATCG 68346  
QY 1018 CACACCAAGCTTCAGAGTGAGGATTCATTTGTACCCAGATTTGGCAATTTACCAA 1077  
Db 68347 CTACACCAAGCTTAAAGAGTGAGGATTCATCTGTACAAAGAAATCACTACATTTACCAG 68406  
QY 1078 GCATTTCAAGGATTAATATTAATGAAAGCTGTAAGTAAACTTGGTGAAGGACACA 1137  
Db 68407 GCATTTCAATTAATATTAATTCCTCGTTATGGTTCTGTAACTAACTCGTGAAGGCCACA 68466  
QY 1138 GGTACTATCATTTGAAAACTTTGACAAATGGGTCTTTCTTCCATTACACC--TTCTTTTATT 1195  
Db 68467 CCCTCTGCTTGAACACACTGGCAGTGAACATCTCGCTTGGATCTGTTTCTCCATAAT 68526  
QY 1196 TTAAGGCTGTGAAT 1211  
Db 68527 TTCAAGGCTTCTGAGT 68542  
Search completed: October 14, 2004, 03:42:41  
Job time : 4901.47 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2004, 00:32:26 ; Search time 482.369 Seconds  
(without alignments)  
11167.204 Million cell updates/sec

Title: US-10-790-589-3

Perfect score: 1268

Sequence: 1 atggggaagtgtc

Scoring table: IDENTITY NUM

Gapop 10.0 ; Gap

4

searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters.

100

```
Minimum DB seq Length: 0
Maximum DB seq Length: 300000000
```

00000000007: 175689 55 11111111

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

```
Database : N_Geneseq_29Jan04.*
1: geneseq1980s.*
2: geneseq1990s.*
3: geneseq2000s.*
4: geneseq2001as.*
5: geneseq2001bs.*
6: geneseq2002s.*
7: geneseq2003as.*
8: geneseq2003bs.*
9: geneseq2003cs.*
10: geneseq2004s.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query %	Score	Match	Length	DB	ID	Description
1	1268	100.0	1268	5	AAD21289	Human deo	
2	1260.4	99.4	1690	5	ABV23663	Human pro	
3	1260.4	99.4	1690	5	ABV23788	Human pro	
4	1140	89.9	1140	6	ABK92136	Prostate	
5	1079.6	85.1	1086	4	AAF60716	Human DNA	
6	673.8	53.1	1652	4	AAF60715	Human DNA	
7	672.8	53.1	1224	5	AAD21288	Murine DN	
8	310.8	24.5	480	8	ACH19337	Mouse deo	
9	155	12.2	265145	9	ADE87477	Human adu	
10	84	6.6	1915	2	AAV29137	Powpox v	
11	82.4	6.5	1575	2	AAV02903	Homo sapi	
12	71	5.6	927	2	AAV29138	Human DNA	
13	60	4.7	60	6	ABN04074	Bos tauru	
14	57	4.5	65	6	ABN29670	Human spl	
15	54.4	4.3	544	7	ACD98178	Rat splic	
16	50.6	4.0	478	8	ABX13974	Human col	
17	46.6	3.7	222	7	ABX55053	Human adu	
18	45	3.5	2000	7	ADA71938	Bovine ES	
19	43.4	3.4	2000	7	ADA71938	Human ES	
20	39	3.1	16998	6	AAD36511	Rice gene	
21	39	3.1	197496	6	ABN85584	Rice gene	
22	38.6	3.0	8759	7	ABZ10109	Human Her	
23	38.6	3.0	8759	7	ABZ10237	Human EGF	
						Haematopo	
						ABZ10237	

Adh5454213	Preteate
Adg841421	Human lym
Ages24034	DNA encod
Aad323094	Human neu
Aa195178	Human neu
Aa119958	Human bre
Aas46344	Tumour su
Adh81251	Signal tra
Adh80239	Human che
Aa125538	Drosophi
Aa132633	Human inm
Adh67077	Human ang
AaA86899	Human inm
AaA86898	Human inm
AaA86897	Human inm
Adh671243	Human bra
AaA161371	Soybean 2
AaA161372	Soybean 2
Adc86916	Human GPC
Adh331189	Signal tra
AaA631086	Human gen
Adh133133	Drosophi
Adh133132	Drosophi

## ALIGNMENTS

RESULT 1	
AAD21289	
ID	AAD21289 standard; cDNA; 1268 BP.
XX	
AC	AAD21289;
XX	
DT	28-JAN-2002 (first entry)
XX	
DE	Human deoxyribonuclease (DNase) II beta cDNA.
XX	
KW	Human; deoxyribonuclease; DNase II beta protein; mucous plug; lung;
XX	cystic fibrosis; pulmonary; chromosome lp22; ss.
OS	Homo sapiens.
XX	
Key	Location/Qualifiers
CDS	46..1119
FT	/*tag= a
FT	/product= "Human deoxyribonuclease (DNase) II beta
FT	protein"
FT	

WO200175082-A1.

11-OCT-2001.

02-APR-2001; 2001WO-US070635

003-APR-2000: 2000TS-00541840

13-MAY-2000; 2000US-00574942.

(DART-) DARTMOUTH COLLEGE.

Bascom AR, Krieser RJ;

REF: Z001-6629/2/16.  
P-PSDB: AAE13015

New cDNA encoding a deoxyribo-

ANA present in the mucous plug

Page 15-16; 21pp; Eng

The invention relates to deoxy-

the mucous plugs in lungs and their viscosity. The present

(f) assessing the prostate cell carcinogenic potential of a compound;

CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker  
yx

Query Match 99.4%; Score 1260.4; DB 5; Length 1690;  
Best Local Similarity 99.9%; Pred. NO. 0;  
Matches 1261; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	7	AAAGTGTCTGCTGTGGCATGAAATAAAATGAAACAGAAAAATCATGCCAGAGCTGCTAAGA	66
Db	7	AAAGTGTCTGCTGTGGCATGAAATAAAATGAAACAGAAAAATCATGCCAGAGCTGCTAAGA	66
QY	67	ACATCCITTTGCTTTGCTCTTCTGGCCTCTTTTGGGCTCTTTGGGCTCTGGGGGCGACCAAAATTTCA	126
Db	67	ACATCCITTTGCTTTGCTCTTCTGGCCTCTTTTGGGCTCTTTGGGCTCTGGGGGCGACCAAAATTTCA	126
QY	127	TGCAGAAATGAAGAGGGAAGACCTGTGACNCTGGTTTACTTTTTATAGATTACCTTAAAGA	186
Db	127	TGCAGAAATGAAGAGGGAAGACCTGTGACNCTGGTTTACTTTTTATAGATTACCTTAAAGA	186
QY	187	CAAAACAAGGAAAGTGGAGAGACTGGGTTAGAGTACCTGTACCTAGACTCTTACAACTAGA	246
Db	187	CAAAACAAGGAAAGTGGAGAGACTGGGTTAGAGTACCTGTACCTAGACTCTTACAACTAGA	246
QY	247	AGCTTGGAGGAAGAGTGAGCAACTAATGAATGACCAAGAGTGTTTTGGGAAGGACAAATTA	306
Db	247	AGCTTGGAGGAAGAGTGAGCAACTAATGAATGACCAAGAGTGTTTTGGGAAGGACAAATTA	306
QY	307	CAACAGCTATATGAAGCATATGCCCTTAAGAGTAAACAACAGAGCTATCTAAATATACAAT	366
Db	307	CAACAGCTATATGAAGCATATGCCCTTAAGAGTAAACAACAGAGCTATCTAAATATACAAT	366
QY	367	GATGGAGTCCCTAAACCTGTGAATTAAGTAGAAAAGTATGGAACAACAAGGTTTACTG	426
Db	367	GATGGAGTCCCTAAACCTGTGAATTAAGTAGAAAAGTATGGAACAACAAGGTTTACTG	426
QY	427	CTGTGGAACAAGAGTTCAAGGTTCTGTGCTGATTCATTCCTCCATCCCTCAGTTTCTCCAAAT	486
Db	427	CTGTGGAACAAGAGTTCAAGGTTCTGTGCTGATTCATTCCTCCATCCCTCAGTTTCTCCAAAT	486
QY	487	CCGGAAGAAGGCTATGATTAATCCACCAAGGAGAGCAAAATGGACAAAGTGGCATCTGC	546
Db	487	CCGGAAGAAGGCTATGATTAATCCACCAAGGAGAGCAAAATGGACAAAGTGGCATCTGC	546
QY	547	ATAACTTTCAAGTACAACAGTAGTAGGCAATAGATTCAGCTCTTGTGTTCTGCAACCCC	606
Db	547	ATAACTTTCAAGTACAACAGTAGTAGGCAATAGATTCAGCTCTTGTGTTCTGCAACCCC	606
QY	607	AAGCTCTATAGCTGCTCCATCCAGACACCTTTTCCAGGAGGCTCAITCAATGCCCCAG	666
Db	607	AAGCTCTATAGCTGCTCCATCCAGACACCTTTTCCAGGAGGCTCAITCAATGCCCCAG	666
QY	667	CTGTGACCAAGGCCCAGCTCATCAGAGATTCCTGGCAGGCTCTCTACCACTTCAGTCTG	726
Db	667	CTGTGACCAAGGCCCAGCTCATCAGAGATTCCTGGCAGGCTCTCTACCACTTCAGTCTG	726
QY	727	GCCCAAGGCAAAAATTCCTCCAATTTGCAAAAGTCGGATTCTTTTCTTGACGACATCTTT	786
Db	727	GCCCAAGGCAAAAATTCCTCCAATTTGCAAAAGTCGGATTCTTTTCTTGACGACATCTTT	786
QY	787	GCAGCTGTGATGCTCAACGGCTGAAGACACACTGTTTAAAGAAACCTGGCAGCGAAA	846
Db	787	GCAGCTGTGATGCTCAACGGCTGAAGACACACTGTTTAAAGAAACCTGGCAGCGAAA	846
QY	847	AGACAAGAGCTTCCCTTCAAACTGCTCCCTTCCATCCATGCTCAAAATATAAGCAAT	906
Db	847	AGACAAGAGCTTCCCTTCAAACTGCTCCCTTCCATCCATGCTCAAAATATAAGCAAT	906
QY	907	AAATATTACGACACTTCTTATTTTCAAGTCTTATCAGATCAAGCAAGTGGTGTATTTCC	966
Db	907	AAATATTACGACACTTCTTATTTTCAAGTCTTATCAGATCAAGCAAGTGGTGTATTTCC	966

QY	967	CAAAAGGGCACCAAAATCGCTGGACATGTATTGGAGACCTAAATCGAGTCCACACAA	1026
Db	967	CAAAAGGGCACCAAAATCGCTGGACATGTATTGGAGACCTAAATCGAGTCCACACAA	1026
QY	1027	GCCTTCAGAAGTGGAGGATTCAATTGTATCCAGAATTGGCAAAATTTACCAGCAATTTCAA	1086
Db	1027	GCCTTCAGAAGTGGAGGATTCAATTGTATCCAGAATTGGCAAAATTTACCAGCAATTTCAA	1086
QY	1087	GSATTTAGTATTATACATGAAAGCTGTAAAGTAAACTTTGGTGAAGGACACAGGTACTATC	1146
Db	1087	GSATTTAGTATTATACATGAAAGCTGTAAAGTAAACTTTGGTGAAGGACACAGGTACTATC	1146
QY	1147	ATTGAAAACCTTTGACAATGGGCTCTTCTTCCATTACACCTTCTTTATATTTTAAAGGCGTG	1206
Db	1147	ATTGAAAACCTTTGACAATGGGCTCTTCTTCCATTACACCTTCTTTATATTTTAAAGGCGTG	1206
QY	1207	TGAATATACTTATTAACCTTGCATATCAAAAATAAAACATATTTCTCTCATGTTTACCATT	1266
Db	1207	TGAATATACTTATTAACCTTGCATATCAAAAATAAAACATATTTCTCTCATGTTTACCATT	1266
QY	1267	TA 1268	
Db	1267	TA 1268	

RESULT 3  
ABV23788

ID ABV23788 standard; cDNA; 1690 BP.

XX  
AC ABV23788:

XX  
DT 16-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 23779.

xx  
kw Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
kw pharmacogenomic marker; gene; ss.

XX  
OS Homo sapiens.XX  
PN WO200160860-A2

XX  
PD 23-AUG-2001.

XX  
PF 20-FEB-2001; 2001WO-US005171.

XX  
PR 17-FEB-2000; 2000US-0183319P.

PR 16-MAR-2000; 2000US-0189862P.  
PR 25-MAY-2000; 2000US-0207454P.  
PR 25-MAY-2000; 2000US-0207454P.

PR 09-JUN-2000; 2000US-0211314P.  
PR 18-JUL-2000; 2000US-0219007P.  
PR 12-DEC-2000; 2000US-0219007P.

PR 13-DEC-2000; 2000US-0255281P.  
XX  
PA /MILITARY: 2000US-0255281P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE; .....

XX  
DR WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.  
vz  
vz

PS Claim 1; Page 4376; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy  
CC of progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;

CC	(e) selecting a composition for inhibiting prostate cancer in a patient;
CC	(f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC	determining whether prostate cancer has metastasized in a patient; (h)
CC	assessing the aggressiveness or indolence of prostate cancer in a patient
CC	; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX	
SEQ	Sequence 1690 BP; 525 A; 369 C; 305 G; 485 T; 0 U; 6 Other;
	Query Match 99.4%; Score 1260.4; DB 5; Length 1690;
	Best Local Similarity 99.9%; Pred. No. 0;
	Matches 1261; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	7 AAAGTGTCTCGCTGTGGCATGAATAAATGAACAGAAAATGATGGCAAGACTGCTAAGA 66
Db	7 AAAGTGTCTCGCTGTGGCATGAATAAATGAACAGAAAATGATGGCAAGACTGCTAAGA 66
QY	67 ACATCCTTTGCTTTGCTCTTTCCTTGGGCTCTTTGGGGTGTCTGGGGGAGCAAACTTCA 126
Db	67 ACATCCTTTGCTTTGCTCTTTCCTTGGGCTCTTTGGGGTGTCTGGGGGAGCAAACTTCA 126
QY	127 TGCAGAAATCAAGAAAGGAAAGCTGTGGACTGTGTTTATATAGTTTACCTTAAAGA 186
Db	127 TGCAGAAATCAAGAAAGGAAAGCTGTGGACTGTGTTTATATAGTTTACCTTAAAGA 186
QY	187 CAAAAAAGAAAGTGGAGAGACTGGGTTAGAGTACCTGTACCTAGACTCTACAACTAGA 246
Db	187 CAAAAAAGAAAGTGGAGAGACTGGGTTAGAGTACCTGTACCTAGACTCTACAACTAGA 246
QY	247 AGCTGGAGGAGAGTGGACACTAATGAATGACACCAAGAGTGTTTGGGAAAGACATTA 306
Db	247 AGCTGGAGGAGAGTGGACACTAATGAATGACACCAAGAGTGTTTGGGAAAGACATTA 306
QY	307 CAACAGCTATATGAAGCATATGCCCTCTAAGAGTAAACAACAGAGCTATCTAATATACAAT 366
Db	307 CAACAGCTATATGAAGCATATGCCCTCTAAGAGTAAACAACAGAGCTATCTAATATACAAT 366
QY	367 GATGGAGTCCCTAAACCTGTGAATTAAGTAAAGTATGACACACCAAGAGTTTACTG 426
Db	367 GATGGAGTCCCTAAACCTGTGAATTAAGTAAAGTATGACACACCAAGAGTTTACTG 426
QY	427 CTGTGGAACAGAGTTCAAGGGTCTGGCTGATTCATTCCATCCCTCAGTTTCCCTCCAATT 486
Db	427 CTGTGGAACAGAGTTCAAGGGTCTGGCTGATTCATTCCATCCCTCAGTTTCCCTCCAATT 486
QY	487 CCGGAAGAAGGCTATGATATTATCCACCCACAGGGAGACGAAATGGACAAAGTGGCATCTGC 546
Db	487 CCGGAAGAAGGCTATGATATTATCCACCCACAGGGAGACGAAATGGACAAAGTGGCATCTGC 546
QY	547 ATAACTTTCAAGTACAAACAGTATGAGGCAATAGATTCTCAGCTCTTGCTGCAACCCC 606
Db	547 ATAACTTTCAAGTACAAACAGTATGAGGCAATAGATTCTCAGCTCTTGCTGCAACCCC 606
QY	607 AACGTCTATAGCTGTCCATCCACAGCACCTTTACACAGGAGCTCATTCACATGCCCCAG 666
Db	607 AACGTCTATAGCTGTCCATCCACAGCACCTTTACACAGGAGCTCATTCACATGCCCCAG 666
QY	667 CTGTGCAACAGGGCCAGCTCATCAGAGATTCTGGCAGGCTCCTCAACAATTCTCAGTGC 726
Db	667 CTGTGCAACAGGGCCAGCTCATCAGAGATTCTGGCAGGCTCCTCAACAATTCTCAGTGC 726
QY	727 GCCCGGGGACAAAATTCCTCCATTGTCAGAGTCGGATTCTTTCTTGACGACATCTTT 786
Db	727 GCCCGGGGACAAAATTCCTCCATTGTCAGAGTCGGATTCTTTCTTGACGACATCTTT 786
QY	787 GCAGCTGATGGCTCAACGGGTGAAGACACACTTGTAAACGAAACCTGGCAGCGAAAA 846
Db	787 GCAGCTGATGGCTCAACGGGTGAAGACACACTTGTAAACGAAACCTGGCAGCGAAAA 846
QY	847 AGCAAGAGGCTTCCTTCAAACGCTCCCTTCCCTTACCATGTCTCAATATAAAGCAATT 906
Db	847 AGCAAGAGGCTTCCTTCAAACGCTCCCTTCCCTTACCATGTCTCAATATAAAGCAATT 906
QY	907 AAATTTACGACACTCTTATTTTCAAGTCTTATCAAGATCAACCCCAAGTGGTGTAATTTCC 966



XX The present invention relates to methods of detecting a prostate cancer-  
 CC associated transcript in a cell from a patient. The method comprises  
 CC contacting a biological sample from the patient with prostate cancer-  
 CC associated polynucleotides (designated PC genes) that selectively  
 CC hybridize to a sequence that is at least 80% identical to them. The  
 CC prostate cancer-associated polynucleotide sequences are differentially  
 CC expressed in prostate tumour tissue or in prostate cancer and are derived  
 CC from the tissues of various organisms such as humans or other mammals  
 CC (e.g. mice, sheep and dogs). The methods of the invention are useful for  
 CC diagnosing and treating prostate cancer in mammals. The prostate cancer-  
 CC associated genes are useful for diagnosing or treating prostate cancer,  
 CC as well as for identifying modulators of prostate cancer or agents that  
 CC inhibit prostate cancer. The nucleic acid sequences are particularly  
 CC useful in gene therapy, as a vaccine or in antisense applications.  
 CC ABK92115-ABK92263 represent prostate cancer-associated polynucleotide  
 CC sequences  
 XX  
 SQ Sequence 1140 BP; 357 A; 258 C; 245 G; 280 T; 0 U; 0 Other;  
 Query Match 89.9%; Score 1140; DB 6; Length 1140;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGGGAAAGTGTCTCTGCTGTCATGAATAAATGAACAGAAATGATGGCAAGACTG 60  
 DB 1 ATGGGAAAGTGTCTCTGCTGTCATGAATAAATGAACAGAAATGATGGCAAGACTG 60  
 QY 61 CTAAGAACATCTTCTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 120  
 DB 61 CTAAGAACATCTTCTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 120  
 QY 121 ATTTTCATGAGAAATGAAGAGGAAAGCTGTGGACTGTGGTGTCTTCTTCTTCTTCTTCT 180  
 DB 121 ATTTTCATGAGAAATGAAGAGGAAAGCTGTGGACTGTGGTGTCTTCTTCTTCTTCTTCT 180  
 QY 181 AAAAGACAAAACAGAAAGTGGAGAGTGGGTGTAGAGTACCTGTACCTAGACTCTACA 240  
 DB 181 AAAAGACAAAACAGAAAGTGGAGAGTGGGTGTAGAGTACCTGTACCTAGACTCTACA 240  
 QY 241 ACTAGAGCTGGAGAGAGTGGAGAGTGGGTGTAGAGTACCTGTACCTAGACTCTACA 300  
 DB 241 ACTAGAGCTGGAGAGAGTGGAGAGTGGGTGTAGAGTACCTGTACCTAGACTCTACA 300  
 QY 301 ACATTACACAGCTATATGAACATATGCTCTTAAGAGTAAACACACAGCTTCTTAATA 360  
 DB 301 ACATTACACAGCTATATGAACATATGCTCTTAAGAGTAAACACACAGCTTCTTAATA 360  
 QY 361 TACAATGATGGAGTCCCTAAACCTGTGAATTAAGTGAAGATATGGACACACCAAGGT 420  
 DB 361 TACAATGATGGAGTCCCTAAACCTGTGAATTAAGTGAAGATATGGACACACCAAGGT 420  
 QY 421 TTACTGCTGTGAACAGAGTTCAGAGGTTCGGCTGATTCATTCCTCCTCAGTTTCTCT 480  
 DB 421 TTACTGCTGTGAACAGAGTTCAGAGGTTCGGCTGATTCATTCCTCCTCAGTTTCTCT 480  
 QY 481 CCAATTCGGAGAGGCTATGATTAATCCACACAGAGGAGCAAAATGGACAAAGTGGC 540  
 DB 481 CCAATTCGGAGAGGCTATGATTAATCCACACAGAGGAGCAAAATGGACAAAGTGGC 540  
 QY 541 ATCTGCTATCTTTCAAGTACAAACAGATGATGAGCAATAGATTTCTGCTGCTGCG 600  
 DB 541 ATCTGCTATCTTTCAAGTACAAACAGATGATGAGCAATAGATTTCTGCTGCTGCG 600  
 QY 601 AACCCACAGCTATAGTCTGCTCCATCCAGGACCTTTCACAGAGCTCATTCACATG 660  
 DB 601 AACCCACAGCTATAGTCTGCTCCATCCAGGACCTTTCACAGAGCTCATTCACATG 660  
 QY 661 CCCAGCTGTGCAACAGGCGCAGCTCATCAGAGATTCCTGACAGCTCTCTACACACTT 720  
 DB 661 CCCAGCTGTGCAACAGGCGCAGCTCATCAGAGATTCCTGACAGCTCTCTACACACTT 720  
 QY 721 CAGTCGGCCCGAGGACAAAAATTCCTCCATTTTGGCAAGTCGGATTCCTTTCTTGAACGAC 780

DB 721 CAGTCGGCCCGAGGACAAAAATTCCTCCATTTTGGCAAGTCGGATTCCTTTCTTGAACGAC 780  
 QY 781 ATCTTTGACGCTGGATGGCTCAACGGCTCAAGACACACACTTTTAAACAGAAACCTGGCAG 840  
 DB 781 ATCTTTGACGCTGGATGGCTCAACGGCTCAAGACACACACTTTTAAACAGAAACCTGGCAG 840  
 QY 841 CGAAAAGACAGAGCTCTCTTCAAACTGCTCCCTTCCCTTACCATGCTCTACATATATAA 900  
 DB 841 CGAAAAGACAGAGCTCTCTTCAAACTGCTCCCTTCCCTTACCATGCTCTACATATATAA 900  
 QY 901 GCAATTAATATATACAGACACTCTTATTTTCAAGTCTTATCAAGATCAGCCAGTGTGT 960  
 DB 901 GCAATTAATATATACAGACACTCTTATTTTCAAGTCTTATCAAGATCAGCCAGTGTGT 960  
 QY 961 ATTTCCCAAAAGGGCCACCAAAAATCGCTGACATGATTTGGAGACCTTAAATCGAGTCCA 1020  
 DB 961 ATTTCCCAAAAGGGCCACCAAAAATCGCTGACATGATTTGGAGACCTTAAATCGAGTCCA 1020  
 QY 1021 CACCAAGCTTTCAGAGTGGAGGATTCATTTTACCCAGAAATTTGCAAAATTTTACCAGCA 1080  
 DB 1021 CACCAAGCTTTCAGAGTGGAGGATTCATTTTACCCAGAAATTTGCAAAATTTTACCAGCA 1080  
 QY 1081 TTTTCAAGGATTAAGTATATATATGAAAGCTGTAAAGTCTGTAAGTAACTTGTGAAAGGACAGGT 1140  
 DB 1081 TTTTCAAGGATTAAGTATATATGAAAGCTGTAAAGTCTGTAAGTAACTTGTGTAAGGACAGGT 1140  
 RESULT 5  
 AAF60716  
 ID AAF60716 standard; cDNA; 1086 BP.  
 XX AAF60716;  
 XX 03-MAY-2001 (first entry)  
 XX Human DNase coding sequence.  
 XX Human; DNase; DLAD; deoxyribonuclease; endonuclease; cystic fibrosis;  
 XX infectious disease; ss.  
 XX Homo sapiens.  
 XX WO200112793-A1.  
 XX 22-FEB-2001.  
 XX 01-MAY-2000; 2000WO-JP002893.  
 XX 17-AUG-1999; 99JP-00230870.  
 XX (TANU/) TANUMA S.  
 XX Tanuma S, Shiokawa D;  
 XX WPI; 2001-218348/22.  
 XX P-PSDB; AAB72417.  
 XX Acidic deoxyribonuclease capable of divalent cation-independent cleavage  
 XX of DNA under acidic even neutral pH and not inhibited by G-actin, useful  
 XX in remedies for cystic fibrosis and for treatment of infectious diseases.  
 XX Claim 16; Page 53-55; 61pp; Japanese.  
 XX The present sequence is the coding sequence of a human deoxyribonuclease  
 XX (DLAD), which is an endonuclease. DLAD is capable of divalent cation-  
 XX independent cleavage of DNA under acidic conditions. DLAD can be used as  
 XX a substitute for DNase I in treating cystic fibrosis, and is useful in  
 XX the prevention and treatment of infectious diseases  
 XX Sequence 1086 BP; 339 A; 250 C; 228 G; 269 T; 0 U; 0 Other;  
 Query Match 85.1%; Score 1079.6; DB 4; Length 1086;

	Best Local Similarity	99.6%; Pred No. 0;	Mismatches	4; Indels	0; Gaps	0;
	Matches 1082; Conservative	0;				
QY	34	ATGAAACAGAAAAATGATGGCAAGACTCTTAAGAACATCCTTTGCTTGTCTTCCTTGCC	93			
Dd	1	ATGAAACAGAAAAATGATGGCAAGACTCTTAAGAACATCCTTTGCTTGTCTTCCTTGCC	60			
QY	94	CTCTTTGGGGTGTGGGGGAGCAACAATTTCATGCAGAAATGAAGAAGGAAAAGCTGTG	153			
Dd	61	CTCTTTGGGGTGTGGGGGAGCAACAATTTCATGCAGAAATGAAGAAGGAAAAGCTGTG	120			
QY	154	GACTGGTTTTACTTTTTATTAAGTTACCCTAAAGACAAAAAAGGAAAGTGGAGACACTGGG	213			
Dd	121	GACTGGTTTTACTTTTTATTAAGTTACCCTAAAGACAAAAAAGGAAAGTGGAGACACTGGG	180			
QY	214	TTAGAGTACCTGTACCTTAGACTCTTACAACTAGAAGCTGGAGAGAGTGGCAACTAATG	273			
Dd	181	TTAGAGTACCTGTACCTTAGACTCTTACAACTAGAAGCTGGAGAGAGTGGCAACTAATG	240			
QY	274	AATCACACCAAGAGTGTTCCTGGGAAGGACATTACAAACAGCTATATGAAGCAATATGCCCTCT	333			
Dd	241	AATGACACCAAGAGTGTTCCTGGGAAGGACATTACAAACAGCTATATGAAGCAATATGCCCTCT	300			
QY	334	AAGAGTAAACAACACAGCCTATCTTAATAFCAATGATGAGTCCCTAAACCTGTGNAATTAC	393			
Dd	301	AAGAGTAAACAACACAGCCTATCTTAATAFCAATGATGAGTCCCTAAACCTGTGNAATTAC	360			
QY	394	AGTAGAAGTATGGACACACCAAAGTTTACTGCTGTGGAACAGAGTTTCAAGGGTTCCTGG	453			
Dd	361	AGCGAAAAGTATGGACACACCAAAGTTTACTGCTGTGGAACAGAGTTTCAAGGGTTCCTGG	420			
QY	454	CTGATTTCATTCATCCCTCAGTTTCCTCCAATTCGGAAGAGGCTATGATTATCCACCCC	513			
Dd	421	CTGATTTCATTCATCCCTCAGTTTCCTCCAATTCGGAAGAGGCTATGATTATCCACCCC	480			
QY	514	ACAGGGAGACGAAATFGGACAAAGTGGCATCTGCATAACTTCAAGTACAAACAGTATGAG	573			
Dd	481	ACAGGGAGACGAAATFGGACAAAGTGGCATCTGCATAACTTCAAGTACAAACAGTATGAG	540			
QY	574	GCAATAGATTCTCAGCTCTTGTCTGTCAACCCCAAGCTCTATAGCTGTCTCCATCCAGCC	633			
Dd	541	GCAATAGATTCTCAGCTCTTGTCTGTCAACCCCAAGCTCTATAGCTGTCTCCATCCAGCC	600			
QY	634	ACCTTTTCCACGAGAGCTCAATTCACATGCCCCAGCTGTGCACACGGGCGAGCTCATCAGAG	693			
Dd	601	ACCTTTTCCACGAGAGCTCAATTCACATGCCCCAGCTGTGCACACGGGCGAGCTCATCAGAG	660			
QY	694	ATTCTCGGAGCTCTCACCAACTTCAGTGGCCCGAGGACAAAAATTCCTCCAAATTT	753			
Dd	661	ATTCTCGGAGCTCTCTCAACACATTCAGTGGCCCGAGGACAAAAATTCCTCCAAATTT	720			
QY	754	GCAAGTGGGATTCTTTTCTTGACGACATCTTTGACGCTGGATGGCTCAACGGCTGAAG	813			
Dd	721	GCAAGTGGGATTCTTTTCTTGATGSCATCTTTGACGCTGGATGGCTCAACGGCTGAAG	780			
QY	814	ACACACTTCTTTAACAGAAAACCTGGCGAGCGAAAAGACAGAGCTTCCCTTCAAACTGCTCC	873			
Dd	781	ACACACTTGTAAACAGAAAACCTGGCGAGCGAAAAGACAGAGCTTCCCTTCAAACTGCTCC	840			
QY	874	CCTTCTTACCATGTCTTACATAATAAAGCAATTTAAATTTATCAGACACTCTTATTTTCAGT	933			
Dd	841	CCTTCTTACCATGTCTTACATAATAAAGCAATTTAAATTTATCAGACACTCTTATTTTCAGT	900			
QY	934	TCTTATCAAGATCACGCCAAGTGGTGTATTTCCAAAAGGGCACCAAAAATTCGCTGACAC	993			
Dd	901	TCTTATCAAGATCACGCCAAGTGGTGTATTTCCAAAAGGGCACCAAAAATTCGCTGACAC	960			
QY	994	TGTTATTTGGAGACCTTAATTCGAGTCCACACCAAGCCTTCAGAGTGGAGATTCAATTTGT	1053			
Dd	961	TGTTATTTGGAGACCTTAATTCGAGTCCACACCAAGCCTTCAGAGTGGAGATTCAATTTGT	1020			
QY	1054	ACCCGAAATTGGCAAAATTTTACCAGCAATTTCAAGGATTAGTATTATCTATGAAGCTGT	1111			

Db	1021	ACCAGAAATGGCAAAATTTACCAAGCATTTTCAAGGATTAGTAGTATTATATGAAAGCTGT	1081
QY	1114	AACTAA 1119	
Db	1081	AACTAA 1086	
RESULT 6			
AAAF60715			
ID	AAAF60715	standard; cDNA; 1652 BP.	
XX	XX		
AC	AAAF60715;		
XX	XX		
DT	03-MAY-2001	(first entry)	
XX	XX		
DE	Murine DNase coding sequence.		
XX	XX		
XX	Murine; DNase; DLAD; deoxyribonuclease; endonuclease; cystic fibrosis;		
KW	infectious disease; ss.		
XX	XX		
OS	Mus musculus.		
XX	XX		
PN	WO200112793-A1.		
XX	XX		
PD	22-FEB-2001.		
XX	XX		
EF	01-MAY-2000; 2000WO-JP002893.		
XX	XX		
PR	17-AUG-1999; 99JP-00230870.		
XX	XX		
PA	(TANU/) TANUMA S.		
XX	XX		
PI	Tanuma S, Shiokawa D;		
XX	XX		
DR	WPI; 2001-218348/22.		
DR	P-PSDB; AAB72416.		
XX	XX		
PT	Acidic deoxyribonuclease capable of divalent cation-independent cleavage		
PT	of DNA under acidic even neutral pH and not inhibited by G-actin, useful		
PT	in remedies for cystic fibrosis and for treatment of infectious diseases.		
XX	XX		
PS	Claim 14; Page 49-51; 61pp; Japanese.		
XX	XX		
CC	The present sequence is the coding sequence of a murine deoxyribonuclease		
CC	(DLAD), which is an endonuclease. DLAD is capable of divalent cation-		
CC	independent cleavage of DNA under acidic conditions. DLAD can be used as		
CC	a substitute for DNase I in treating cystic fibrosis, and is useful in		
CC	the prevention and treatment of infectious diseases		
XX	XX		
SQ	Sequence 1652 BP; 494 A; 410 C; 339 G; 409 T; 0 U; 0 Other;		
Query Match 53.1%; Score 673.8; DB 4; Length 1652;			
Best Local Similarity 75.1%; Pred. No. 1.3e-194;			
Matches 884; Conservative 0; Mismatches 282; Indels 11; Gaps 3			
QY	37	AAACAGAAATCATGGCAGACTGCTTAAGACATCCTTTGCTTGTCTCTTCCTTGGCCTC 96	
Db	204	AGATAGAAATGACAGCAAGCCTCTAAGACAGTTCTTTTGTCTCTTTTGGCCTC 263	
QY	97	TTTGGGGTCTGGGGCAGCAACAATTCATGCAGAAATCAAGAGGAAAGCTGCGAC 156	
Db	264	TCTGGGCTCTGGGGACACAGAAATCTCATGCAGAAATGAATATGTTGAAGCTGTGGAC 323	
QY	157	TGCTTTACTTTTTTATAAGTTACTCTAAAGACAAAACAAGGAAAGTGCAGAGACTGGTTA 216	
Db	324	TGCTTTACTTTTATAAGTTATCCCAAAGGACTAGCAGGCAAGTCAAGAGGGGGGCTG 383	
QY	217	GAGTACCTGTACTAGACTCTCAACTAGAGCTGGAGGAGACTGAGCAACTAATCAAT 276	
Db	384	CAGTACCTGTACTGTGACTCCACAAGACAAACCTGGAACAGAGCCTCTACCTGATTAC 443	
QY	277	GACACCAAGAGTCTTTTGGGAGGACATTACACAGCTATATCAAGCATATGCTCTTAAG 330	



Fri Oct 15 10:01:04 2004

us-10-790-589-3.rng

Db 448 ATACACTGTGTTCCCAAGTTTCCCCAGTTC-----ATGGCTATGAGTACCAACCTCG 501  
QY 517 GCGAGACGAAATGGACAAAGTGGCATCTGCTAATACTTTCAAGTACAAACAGATATGAGGCA 576  
Db 502 GCGAGCGGATATGGACAAACCGGCATCTGCATCACTTTCGGATACAGCCAGTTCGAGGAA 561  
QY 577 ATAGATTCAGCTCTTGTGTTGCTGCAACCCCAAGCTATAGCTGCTCCATCCAGCCACC 636  
Db 562 ATAGATTTTCAGCTCTTGTGTTGCTGCAACCCCAAGCTATAGCTGCTCCATCCAGCCACC 621  
QY 637 TTTTCCAGGAGCTCATTCACATGCCCCCAGCTGTGACACAGGCCAGCTCATCAGAGATT 696  
Db 622 TTTTCACTGGAACCTTATCTACATGCCCCCGGATGTGCGCACTCCAGTTCCTTAAAGATC 681  
QY 697 CTGCGAGGCTCTCCACCACTTCAGTCTGCGCCCGAGGACAAATAATCTCCATTTTGGCA 756  
Db 682 CTTGTCCGGTACCTCGTGAACCTGCACTACGCCAGGGTCTAACTTCGTCCATTTTGGCA 741  
QY 757 AAGTGGATTCCTTTCTTGAGCAGCATCTTTGACGCTGGATGCTCAACGCTGAGGACA 816  
Db 742 AAATCAAGTTTATATCTGATGATCATCTTTACAGGATGGATAGCTCAAAAGTTGAGACA 801  
QY 817 CACTGTGTAACAGAACTGCGACGAGAAAGACAGAGCTTCCTTCAAACTGCTCCCTT 876  
Db 802 CATTTGTTAGCAAACTGCGAGAAAGAGAAACAGAGCTTCCTTCAAACTGCTCCCTG 861  
QY 877 CTTTACCATGCTCTACATATAAAGCAATTAATATACGACACTCTTATTTTCACTTCT 936  
Db 862 CTTTACCATGCTCTACATATAAAGCAATTAATATACGACACTCTTATTTTCACTTCT 921  
QY 937 TATCAAGATCAGCAAGTGTGTATTTCCCAAAGGCGACCAAAATCGCTGACATGT 996  
Db 922 CGCAAGACCATTCCTCAATGTTGTGTTTCCATTAAGGCTCCCGCAATCGCTGACCTGC 981  
QY 997 ATTGAGACCTAAATCGAGTCCACCAAGCCCTTCAGAGTGGAGGATTCATTTGTACC 1056  
Db 982 ATTGAGACCTAAATCGAGTCCACCAAGCCCTTCAGAGTGGAGGATTCATTTGTACA 1041  
QY 1057 CAGAAATGGCAAAATTTACCAAGCATTTCAAGGATATGATATATGATGAAGCTGTAAAG 1116  
Db 1042 AAGAATCACTACATTTACAGGCAATTCATTAATATATCTCCGTTATGGGTTCTGTAAAG 1101  
QY 1117 TAAACTTGGTGAAGGACACAGTACTATCATTTGAAACCTTGCAATGGGTCTTCTTCC 1176  
Db 1102 TAAACTCGTGAAGGCGCACCCCTCTGCTCTTGTGAAACACTGGCACTGGAACTCTCGC 1161  
QY 1177 ATTACACTTCTTTATATTTTA 1198  
Db 1162 CTGGATCTGTCTCCATAATA 1183

RESULT 8  
ACH19937  
ID ACH19937 standard; cDNA; 480 BP.  
AC ACH19937;  
XX  
XX  
XX 13-OCT-2003 (first entry)  
XX Human adult lung cDNA #940.  
XX  
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
KW genome mapping; biodiversity; genetic disorder.  
XX Homo sapiens.  
XX  
XX US2003073623-A1.  
XX  
XX 17-APR-2003.  
XX  
XX 30-JUL-2001; 2001US-00918995.  
XX  
XX 30-JUL-2001; 2001US-00918995.  
PR

XX (DRMA/) DRMANAC R T.  
PA (LABA/) LABAT I.  
PA (STAC/) STACHE-CRAIN B.  
PA (DICK/) DICKSON M C.  
PA (JONE/) JONES L W.  
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
XX WPI; 2003-615964/58.  
XX New polynucleotide sequences obtained from various cDNA libraries. useful  
PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
PT mapping, in the recombinant production of protein, or in generating  
PT antisense DNA or RNA.  
XX Claim 1; SEQ ID NO 7149; 44pp; English.  
XX The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SBH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: the sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030073623  
XX  
SQ Sequence 480 BP; 134 A; 120 C; 106 G; 118 T; 0 U; 2 Other;

Query Match 24.5%; Score 310.8; DB 8; Length 480;  
Best Local Similarity 99.4%; Pred. No. 4.9e-84;  
Matches 312; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 335 AGAGTAAACACACAGCCCTATCTAATATACATGATGAGTCCCTAAACCTGTGAATTACA 394  
Db 166 AGAGTAAACACACAGCCCTATCTAATATACATGATGAGTCCCTAAACCTGTGAATTACA 225  
QY 395 GTAGAAATATGGACACACCAAGGTTTACTGCTGTGGAACAGAGTTTCAAGGTTCTGGC 454  
Db 226 GCAGAAATATGGACACACCAAGGTTTACTGCTGTGGAACAGAGTTTCAAGGTTCTGGC 285  
QY 455 TGATTTCATTCCTCCTCAGTTCCTCCAAATTCGGAAGAGGCTATGATTATCCACCA 514  
Db 286 TGATTTCATTCCTCCTCAGTTCCTCCAAATTCGGAAGAGGCTATGATTATCCACCA 345  
QY 515 CAGGAGACGAAATGGACAAAGTGGCATCTGCATACCTTCAAGTACACAGATGATGAGG 574  
Db 346 CAGGAGACGAAATGGACAAAGTGGCATCTGCATACCTTCAAGTACACAGATGATGAGG 405  
QY 575 CAATAGATTCCTCAGCTCTTGTGCTGCAACCCCAAGCTATAGCTGCTCCATCCAGGCA 634  
Db 406 CAATAGATTCCTCAGCTCTTGTGCTGCAACCCCAAGCTATAGCTGCTCCATCCAGGCA 465  
QY 635 CCTTTTACCAGGAG 648  
Db 466 CCTTTTACCAGGAG 479

RESULT 9  
ADE87477/c  
ID ADE87477 standard; DNA; 266145 BP.  
XX  
XX ADE87477;  
XX

DT	29-JAN-2004	(first entry)	
XX			
DE	Fowlpox virus genome DNA.		
XX			
KW	fowlpox virus; FPV; virucide; tuberculostatic; protozoacide; antipyrctic;		
KW	cytostatic; hepatotropic; antibacterial; vaccine; malaria; tuberculosis;		
KW	East Coast fever; avipox virus; influenza; hepatitis;		
KW	human papilloma virus; tumour; leishmaniasis; listeriosis; theileria;		
XX	gene; ds.		
OS	Fowlpox virus.		
XX			
EN	WO2003047617-A2.		
XX			
PD	12-JUN-2003.		
XX			
PF	02-DEC-2002; 2002WO-GB005411.		
XX			
PR	30-NOV-2001; 2001GB-00028733.		
XX			
PR	30-NOV-2001; 2001US-0334649P.		
XX			
FA	(ISIS-) ISIS INNOVATION LTD.		
XX			
PI	Laidlaw S, Skinner M, Hill A, Gilbert S, Anderson R;		
XX			
DR	WPI; 2003-513700/48.		
XX			
PT	Treating and/or preventing e.g. malaria or tuberculosis, or eliciting an		
PT	immune response, comprises administering a priming composition and a		
PT	boosting composition containing a non-replicating viral vector in either		
PT	order.		
XX			
P8	Claim 30; SEQ ID NO 1; 302pp; English.		
XX			
CC	The invention relates to a fowlpox virus (FPV) genome which has		
CC	modifications in one or more wild-type FPV genes. The invention further		
CC	relates to a novel method for treating and/or preventing a disease in a		
CC	subject comprising administering two compositions, each containing a non-		
CC	replicating viral vector. At least one of the compositions comprises a non-		
CC	poxvirus vector derived from a fowlpox virus. The novel compositions have		
CC	the following activities: virucide, tuberculostatic, protozoacide,		
CC	antipyrctic, cytostatic, hepatotropic, and antibacterial. The non-		
CC	replicating viral vector is useful in a vaccine for an animal,		
CC	particularly a mammal such as a primate, specifically human. The priming		
CC	or boosting composition, or the kit is useful for manufacturing a		
CC	medicament for treating and/or preventing a disease which is, or results		
CC	from, a chronic infection such as malaria, tuberculosis or East Coast		
CC	fever, or for eliciting a T-cell immune response in a subject. Non-		
CC	cultured CEF cells are useful for growing an avipox virus, such as		
CC	fowlpox virus. The method or the vaccine may further be used to treat or		
CC	prevent influenza, hepatitis, human papilloma virus and other viral		
CC	infections, malignancies such as tumours, leishmaniasis, listeriosis, and		
CC	theileria. This polynucleotide sequence represents the DNA of the fowlpox		
XX	virus genome of the invention.		
SQ	Sequence 266145 BP; 92375 A; 41026 C; 40846 G; 91897 T; 0 U; 1 Other;		
	Query Match 12.2%; Score 155; DB 9; Length 266145;		
	Best Local Similarity 52.3%; Pred. No. 6e-35;		
	Matches 528; Conservative 0; Mismatches 445; Indels 36; Gaps 7;		
QY	114 AGCAACAATTTCATGTCAGAAATGAGAAAGTGGAGACTGGTTTACCTTTTATAAA 173		
DB	38784 AGTTAAGTAAATPTGTTAAATGAGAGCGAAGTAGACTGGTATTTTGGTATACAA 38725		
QY	174 GTTACCTTAAAGACAAAACAAGGAAGTGGAGAGCTGGTTAGAGTACCTGACTAG 233		
DB	38724 GCTCCCTAGCTCCAAAATTTAGGCACCTAAAGGTAAT-----GAATATTATACATAGA 38671		
QY	234 CTCTACAACTAGAAGCTGGAGGAAGAGTGGAGCAACTAATGAATGACACCAAGAGTGTTT 293		
DB	38670 TAGTAATAATCTTAATGGAACGAGGTAAAGTACTCTATTAACCTGATACAGTAAAT 38611		



FT CDS /note= "Region indicated in specification"  
 FT 86.1168  
 FT /\*tag= b  
 FT /product= "DNase II"  
 FT sig\_peptide 86.133  
 FT /\*tag= c  
 FT mat\_peptide 134.1165  
 FT /\*tag= d  
 FT /product= "DNase II"

XX WO9740134-A2.  
 XX 30-OCT-1997.  
 XX 23-APR-1997; 97WO-US006664.  
 XX 25-APR-1996; 96US-00639294.  
 XX (GETH ) GENENTECH INC.  
 XX Baker KP, Baron WF;  
 XX WPI; 1997-535820/49.  
 XX P-PSDB; AAW31495.  
 XX Human DNase II and related nucleic acids - useful in protein production  
 XX e.g. for therapeutic use to treat systemic lupus erythematosus and  
 XX pulmonary diseases e.g. cystic fibrosis.  
 XX Claim 1; Fig 1; 30pp; English.  
 XX This sequence encodes a novel human deoxyribonuclease, DNase II. This  
 XX protein is useful to reduce the viscoelasticity of DNA-containing  
 XX material, e.g. mucus, and used to treat patients with pulmonary diseases  
 XX or disorders e.g. cystic fibrosis, or with systemic lupus erythematosus.  
 XX It may also improve antibiotic efficacy in the treatment of abscesses,  
 XX infected lesions etc., provide treatment in non-infected conditions in  
 XX which there is an accumulation of cellular DNA debris, e.g.  
 XX pyelonephritis and be used to degrade DNA in biological samples or in  
 XX diagnostic assays. The nucleic acids and vectors can be used for in vivo  
 XX or ex vivo gene therapy, and antisense oligonucleotides can be produced  
 XX from the nucleic acids which can bind to and prevent expression of  
 XX nucleic acid within cells. The nucleic acids (or a portion) can also be  
 XX used for hybridisation assays for nucleic acids encoding human DNase II  
 XX in a sample, or to identify and isolate nucleic acids sharing substantial  
 XX sequence identity (e.g. encoding naturally-occurring allelic variants of  
 XX human DNase II). The antibodies can be used to detect and measure human  
 XX DNase II in tissues or clinical samples, and in the purification of human  
 XX DNase II. The nucleic acids enable production of human DNase II by  
 XX recombinant DNA methods in quantities sufficient for clinical use, not  
 XX previously possible  
 XX Sequence 1575 BP; 358 A; 465 C; 450 G; 302 T; 0 U; 0 Other;  
 XX  
 XX Query Match 6.5%; Score 82.4; DB 2; Length 1575;  
 XX Best Local Similarity 47.4%; Pred. NO. 5.3e-14;  
 XX Matches 470; Conservative 0; Mismatches 486; Indels 36; Gaps 6;  
 QY 142 GGGAAAGCTGTGGAAGTCTTTTACTTTTATAGTT---ACCTTAAGACAAACAGGAA 198  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 155 GGGCAGCTGTAGACTGTGTTCTGTGTCTACAGTGTCCAGCTCTTAGAGGGTCCGGGAG 214  
 QY 199 AGTGAGAGACTGGTGTAGTACTCTGACCTAGACTCTACACTAGAGCTGGAGGAG 258  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 215 GCGGCGCAGAGGGCTCAGTACAGTATCTGACGAGAGCTCGGAGGCTGGGGAG 274  
 QY 259 AGTGAGCACTAATGATGACACCAAGAGTGTGTTGGGAGGACATTACACAGCTATAT 318  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 275 GGCAGGGCACTCATCAACAGCCGAGGGGCGCTGGGCGGAGCCCTGACCCCTGTAC 334  
 QY 319 GAAGCATATGCTCTAAGAGTAAACACAGCCCTATCTAATATACATGATGAGTCCCT 378  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 335 CGG-----AGCAACACGAGCAGTCCCTCTCTGTCTCTACATGACCAACCGCCT 385

QY 379 AAACC-----TGTGAATTACAGTAGAAGTATGACACACAAAGGTTTACTGCTGG 432  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 386 CAACCCAGCAAGGCTCAGGACTCTTCATGCTGGGACACGAAGGTTCTGCTCTCTT 445  
 QY 433 AACAGAGTTCAAGGGTCTGCTGCTGATTCATTCCATCCCTCAGTTCTCTCCAAATTCGGAA 492  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 446 GACCACGATGGGGCTTCTGGCTGTGTCACAGTGTACCTAACTTCCCTCCACCGGCTCC 505  
 QY 493 ---GAAGGCTATGATTATCCACCCACAGGAGAGCAATGACAAAGTGGGATCTGATA 549  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 506 TCTGCTCATACAGCTGGCTCTATAGCGCTGTACCTACGGGAGAGCCCTGCTCTGTGTG 565  
 QY 550 ACTTTCAAGTACACCAAGTATGAGGCAATAGATTCTCAGCTCTTGGTCTGCAACCCCAAC 609  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 566 TCTTTTCCCTTCTGCTCAGTTCTGAGATGGCAAGCAGCTGACCTACACTACCTACCTCCCTGG 625  
 QY 610 GTCTATAGCTGCTCCATCCAGCACCCTTTCAACAGGAGCTCATTCACATGCCCCAGCTG 669  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 626 GTCTATAACTACCAAGCTTGGAGGATCTTTGCCAGGAATTCCTCGACTTTGGAGATGTG 685  
 QY 670 TGCACCAAGGCGCAGCTCATCAGAGATTCTTGGCAGGCTCTCACCACACTTCAGTCGGCC 729  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 686 GTCAAGGGCCACACAGTTAGCCAGAACCTCTGGAACAGCAGCATCACACTCACATCCACAG 745  
 QY 730 CAGGGAACAAATTCCTCCATTTTGCAGTCCGATTTCTTCTTGACGACATCTTTGCA 789  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 746 GCGGGGCTGTGTTTCCAGAGCTTTGCCAAGTTTCAGCAATTTGGAGATGACCTGTACTCC 805  
 QY 790 GCCTGATGCTCAACGGCTGAAGACACACTTGTTAACAGAACCTGGCAGCGGAAAGA 849  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 806 GGCTGTTGGCAGCAGCCCTTGTACCAACCTGAGTCCAGTCTGGCACAACACTGTA 865  
 QY 850 CAAGAGCTTCTTCAAACTGCTCCCTTCTTACCATGTCTCAATATATAAAGCAATATA 909  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 866 GGCACTCTCCCTCTAACTGCTCGGATCTGGCAGTTCTGAATGTGAACCATAGCT 925  
 QY 910 TTATACAGCA-----CTCTATTTCAGTCTTATCAAGATCAGCCCAAGTGTGTTATT 963  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 926 TTCCCTGGACCAAGCGGCGCAAGTTCAACAGCAGAGGACCACTCCAAATTTGGTGGT 985  
 QY 964 TCCCAAAAGGCGACCAAAATCGCTGACATGATTGGAGACCTAAATCGGAGTCCACAC 1023  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 986 TCCCAAAAGGCGC-----CTGGACCTCGTGGGTGACATGAATCGGAACCAAGGA 1036  
 QY 1024 CAAGCCTTCAAGAGTGGAGGATTCATTGTTACCCAGAAATTCGCAATTTACCAAGATTT 1083  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 1037 GAGGAGCAACGGGTGGGGGACACTGTGTCGCCAGCTGCCAGCCCTCTGGAAAGCCTTC 1096  
 QY 1084 CAAGATTAGTATTATATCTATGAAAGCTGTAA 1115  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 1097 CAGCGCTGTGTAAGAACTACCAAGCCCTGTAA 1128

RESULT 12  
 AAV29138  
 ID AAV29138 standard; cDNA; 927 BP.  
 XX AC AAV29138;  
 XX DT 11-SEP-1998 (first entry)  
 XX Bos taurus DNase II gene.  
 DE DNase II; deoxyribonuclease II; cystic fibrosis; DNA digestion;  
 XX controlled cell death; apoptosis; metamorphosis; cell turnover;  
 KW neurodegeneration; AIDS; acquired immune deficiency syndrome; cancer;  
 KW autoimmune disease; diagnosis; anticancer agents;  
 KW enhanced chromosomal rearrangement; chromosome instability; ss.  
 XX Bos taurus.  
 XX OS  
 XX Key Location/Qualifiers  
 FH



[illegible]





XX

ps Example 9; Page 934; 959pp; English.

XX The invention describes a method of determining open reading frames in  
XX the genome of organism, comprising contacting mRNA from cell of organism  
CC with a single oligonucleotide primer (I) at low stringency, preparing  
CC single-stranded cDNA by reverse transcribing mRNA with (I), amplifying  
CC cDNA, sequencing the product, and repeating the contacting, preparing  
CC and amplifying steps with different primers and sequencing resulting  
CC nucleic acids. The method is useful for: determining that a known  
CC nucleotide sequence from a genome of an organism corresponds to a  
CC nucleotide sequence of an open reading frame; for preparing a contig,  
CC nucleic acid molecule from a genome of an organism; and for sequencing  
CC all or part of a genome of an organism. mRNA is obtained from mammalian  
CC or human cell which is associated with a pathological condition e.g. a  
CC colon cancer or breast cancer cell. The method is useful for analyses of  
CC populations of subjects and can be used to carry out genetic analyses of  
CC large or small populations. further, it can be used to study living  
CC systems to determine if, e.g. there have been genetic shifts which render  
CC an individual or population more or less likely to be afflicted with  
CC diseases such as cancer, to determine antibiotic resistance or non-  
CC tolerance, and so forth. The method can also be used in the study of  
CC congenital diseases, and the risk of affliction to a fetus, as well as  
CC the study of whether the conditions are likely to be passed to offspring  
CC through ova or sperm. The analyses for pathological conditions can be  
CC carried out in all animals, plants, birds, fish, etc. Using this method,  
CC in the area of agriculture, for example the genomes of food crops can be  
CC studied to determine if resistance genes are present, defects in plant  
CC genomes can also be studied in this way. Similarly, the method permits  
CC determination of the pathogens which integrate into the genome, such as  
CC retroviruses and other integrating viruses such as influenza virus, have  
CC undergone shifts or mutations, which may require different approaches to  
CC therapy. This method is also applied to eukaryotic pathogens, such as  
CC trypanosomes, different types of Plasmodium, etc. The method essentially  
CC eliminates sequencing of non-coding portions. This sequence represents a  
CC polynucleotide isolated from human colon cancer cell cDNA library  
XX

SQ Sequence 544 BP; 114 A; 175 C; 133 G; 122 T; 0 U; 0 Other;

Query Match 4.3%; Score 54.4; DB 7; Length 544;

Best Local Similarity 47.5%; Pred. No. 1.1e-05;

Matches 226; Conservative 0; Mismatches 246; Indels 4; Gaps 2;

QY	400	AAGTATGACACACCAAGGTTTACTGCTGTGGACAGATTCAAGGGTTCTGGCTGATT	459
DB	57	ATGCGTGGGCACACGAGGGTGTCTCTCTTTGACACGATGGGGGCTTTGGTGTGTC	116
QY	460	CATTCCATCCCTCAGTTTCCTCCAA---TTCGGGAAGAGGCTATGATTATCCACCCACA	516
DB	117	CACAGTGATCTACTTCTCCCTCCCGGCTCTCTGTGCATACAGCTGGGCTCATAGC	176
QY	517	GGGAGACGAATGGAACAAAGTGGCATCTGCATACTTTCAAGTACAACAGTATGAGCA	576
DB	177	GCTCTACTCTAGGGCAGACCCCTGCTCTGTGTGTCTTTTCCCTCGCTCAGTTCTCGAAG	236
QY	577	ATAGATTCTCAGCTTGTGTGAAACCCACACCTCTAGCTGTCCATCCAGCCACC	636
DB	237	ATGGGCAAGCAGCTACCTACCTACCCCTGGGTCTATTACTACAGCTGAAGGATC	296
QY	637	TTTTCACAGGAGCTCATTCATCATGCCCCAGCTGTGCACAGGGCCAGCTCATCAGAGATT	696
DB	297	TTTGCCCAAGAAATCCCCGACTTGGAGATGTGGTCAAGGGCCACCACGCTTAGCCAGAA	356
QY	697	CTGGCAGGCTCTCTACCACTTCATGTCGGCCCGAGGACAAAATTCCTCCATTTTGA	756
DB	357	CCCTGGAACAGCAGCATCACACTCACTCCACCGCGGGCTGTT-TTCCAGAGCTTTGCC	415
QY	757	AAGTCGATCTTTTCTTTCAGCAGCATCTTTTGAGCTGTGATGGCTCAACGGCTGAAGACA	816
DB	416	AAGTTACAGCAAAATTTGGAGATGACCTGTACTCCGGCTGGTTGGCAGCAGCCCTTGTACC	475
QY	817	CACCTTGTTTACAGAAACCTGGCAGCGAAAAGACAAGAGCTTCCTTCAAACTGCTC	872

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2004, 00:43:42 ; Search time 106.854 Seconds  
(without alignments)  
6585.417 Million cell updates/sec

Title: US-10-790-589-3

Perfect score: 1268

Sequence: 1 atggggaaagtgtcctgctg.....tctctcatgtttaccattta 1368

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA: \*  
1: /cgn2\_6/prodata/2/ina/5A COMB.seq.\*  
2: /cgn2\_6/prodata/2/ina/5B COMB.seq.\*  
3: /cgn2\_6/prodata/2/ina/6A COMB.seq.\*  
4: /cgn2\_6/prodata/2/ina/6B COMB.seq.\*  
5: /cgn2\_6/prodata/2/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/prodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	1268	100.0	1268	4	US-09-574-942-3
2	1079.6	85.1	1086	4	US-09-807-784B-4
3	673.8	53.1	1652	4	US-09-807-784B-2
4	672.8	53.1	1224	4	US-09-574-942-1
5	84	6.6	1915	3	US-09-147-915-1
6	82.4	6.5	1575	3	US-08-639-294-1
7	82.4	6.5	1575	3	US-09-861-034B-1
8	71	5.6	927	3	US-09-147-915-2
9	51.8	4.1	7218	1	US-08-232-463-14
10	40.2	3.2	505	4	US-09-621-976-15639
11	39.4	3.1	832	4	US-09-621-976-2813
12	39	3.1	169998	4	US-09-676-610B-24
13	39	3.1	197496	4	US-09-877-177A-10
14	34.6	2.7	1509	4	US-09-134-000C-3209
15	34.6	2.7	3321	1	US-08-484-438-5
16	34.6	2.7	5484	3	US-09-632-580A-3
17	34.6	2.7	5501	1	US-08-484-438-1
18	34.4	2.7	364	4	US-09-621-976-17202
19	34.4	2.7	1830121	4	US-09-557-884-1
20	34.4	2.7	1830121	4	US-09-643-990A-1
21	34.2	2.7	8779	4	US-08-956-171E-58
22	34	2.7	3887	4	US-09-976-594-660
23	33.8	2.7	551	4	US-09-389-681-250
24	33.8	2.7	551	4	US-09-620-405B-250
25	33.8	2.7	551	4	US-09-339-338-250
26	33.8	2.7	551	4	US-09-433-826B-250
27	33.8	2.7	551	4	US-09-604-287A-250

c	28	33.8	2.7	551	4	US-09-834-759-250	Sequence 250, Appl
	29	33.6	2.6	918	4	US-09-134-000C-1	Sequence 1, Appli
	30	33.4	2.6	1758	4	US-09-134-000C-2540	Sequence 2540, Ap
c	31	33.4	2.6	640681	4	US-09-790-988-1	Sequence 1, Appli
c	32	33.4	2.6	1664976	4	US-08-916-421B-1	Sequence 1, Appli
c	33	33.2	2.6	5427	3	US-09-009-913-2	Sequence 2, Appli
c	34	33.2	2.6	5510	3	US-09-009-913-3	Sequence 3, Appli
	35	33.2	2.6	5667	3	US-09-009-913-4	Sequence 4, Appli
	36	33	2.6	429	4	US-09-134-000C-214	Sequence 214, Appl
c	37	33	2.6	8607	4	US-10-204-708-72	Sequence 72, Appli
c	38	32.8	2.6	58407	4	US-08-916-421B-2	Sequence 150, Appl
c	39	32.6	2.6	2989	4	US-08-956-171E-150	Sequence 1, Appli
	40	32.4	2.6	708	4	US-08-947-014-1	Sequence 1, Appli
	41	32.4	2.6	708	4	US-09-490-011-1	Sequence 1, Appli
c	42	32.4	2.6	6049	1	US-08-471-033-1	Sequence 1, Appli
c	43	32.4	2.6	6049	2	US-08-471-044-1	Sequence 1, Appli
c	44	32.4	2.6	6049	2	US-08-463-483A-1	Sequence 1, Appli
c	45	32.4	2.6	6049	2	US-08-471-046A-1	Sequence 1, Appli

## ALIGNMENTS

### RESULT 1

US-09-574-942-3  
; Sequence 3, Application US/09574942  
; Patent No. 6358723  
; GENERAL INFORMATION:  
; APPLICANT: Eastman, Alan R.  
; APPLICANT: Krieser, Ronald J.  
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS  
; FILE REFERENCE: DC-0137  
; CURRENT APPLICATION NUMBER: US/09/574,942  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: 09/541,840  
; PRIOR FILING DATE: 2000-04-03  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 1268  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-574-942-3

Query Match	100.0%	Score 1268;	DB 4;	Length 1268;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1268;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	ATGGGGAAAGTCTCTGCTGCTGGCATGAATAATGAACAGAAATGATGGCAAGCTG	60	
Db	1	ATGGGGAAAGTCTCTGCTGCTGGCATGAATAATGAACAGAAATGATGGCAAGCTG	60	
QY	61	CTAAGACATCTTGTCTTGTCTTCTTGGCTCTTGGGCTGCTGGGCGAGCAACA	120	
Db	61	CTAAGACATCTTGTCTTGTCTTCTTGGCTCTTGGGCTGCTGGGCGAGCAACA	120	
QY	121	ATTTCATGCAAAATGAAGAGGAAAGCTGTGGACTGGTTTACTTTTATAGTTACTT	180	
Db	121	ATTTCATGCAAAATGAAGAGGAAAGCTGTGGACTGGTTTACTTTTATAGTTACTT	180	
QY	181	AAAGACAAAACAAAGAAAGTGGAGAGCTGGTTAGTAGTCTGTACTAGACTTACA	240	
Db	181	AAAGACAAAACAAAGAAAGTGGAGAGCTGGTTAGTAGTCTGTACTAGACTTACA	240	
QY	241	ACTAGAAGCTGGAGAGAGTGAAGCACTAATGAATGACCAAGAGTGTGGGAAGG	300	
Db	241	ACTAGAAGCTGGAGAGAGTGAAGCACTAATGAATGACCAAGAGTGTGGGAAGG	300	
QY	301	ACATTACACAGCTATATGAAGCATATGCCCTCTAAGAGTAAACACACAGCCCTATCTAATA	360	
Db	301	ACATTACACAGCTATATGAAGCATATGCCCTCTAAGAGTAAACACACAGCCCTATCTAATA	360	
QY	361	TACATGATGAGTCCCTTAAACCTGTGAATTACAGTAGAAGATGGACACACCAAGGT	420	



Qy	397	AGAAAGTATGGACACACAAAGGTTTACTGCTGTGGAAACAGAGTTTCAAGGGTTCTGGCTG	456
Db	561	AGA CAGTATGGACATGCGAAAGGTTCTGCTGGTATGGAACAGAACGCGAGGGTTCTGGCTG	620
Qy	457	ATTCAATCCATCCCTCAGTTTCCCTCCAAATTCGGGAAGAGCTATGATTATTCACCCACA	516
Db	621	ATACACTCTGTTCCCAAGTTTCCCCAGTTC-----ATGGCTATGAGTACCCAACTTCG	674
Qy	517	GGGAGACGAATGACAAAGTGGCAATCTGCATAACTTTCAAGTACAAACGAGTATGAGGCA	576
Db	675	GGGAGGCGATATGACACAAACGGGCATCTGCATCACTTTCGGATACAGCAGTTTGAGGAA	734
Qy	577	ATAGATTCTCAGCTCTTGCTGTGCAACCCCAACGTCGTATAGTCTCCATCCCGACCAACC	636
Db	735	ATAGATTTTCAGCTCTTGGTCTTCAACCAACAAACATCTACAGCTGCTTCAATTCGAAGCACC	794
Qy	637	TTTCAACAGGAGCTCATTTCAATGCCCCAGCTGTGCACCAAGGGCCAGCTCATCAGAGATT	696
Db	795	TTTCACTGGAAACTTATCTACATGCCCGGATGTGTGCCAACTCCAGTTCCTTAAAGATC	854
Qy	697	CTTGGCAGGCTCTTCACCAACACTTCAGTTCGGCCGACAGGCAAAAATTCCTCCATTTTGCA	756
Db	855	CCGTGTCGGTACTCTCGTGAACATGCATCAGCCAGGCTTAAACTTCGTCCATTTTGCA	914
Qy	757	AAGTCGGATTCTTTCTTTGACGACATCTTTTGACGCTTGGATGCTCAAACGGCTGAAGACA	816
Db	915	AAATCAAGTTTTTATCTGATGATCACTTTTACAGGATGATAGTACTCAAAAGTTGAAGACA	974
Qy	817	CACTTGTTTAAAGAAACCTCGCAGCGGAAAGACAAGAGCTTCTCTTCAAACTGCTCCCTT	876
Db	975	CAITTTGTTAGCAAAACCTGGCAGAAAGAAACAAGAGCTTCCCTTCAAACTGTTCCCTG	1034
Qy	877	CCTTTACCATCTCTACAAATAAAAGCAATTAATTTATCAGACACTCTTATTTTCAGTTCT	936
Db	1035	CCTTTACCATCTCTACAAACATCAAGTCCATTTGGGGTAACTTCCAAGTCTTACTTCAGTTCT	1094
Qy	937	TATCAAGATCACGCCAAGTGGTGTATTTCCAAAAGGGCACCAAAAATCGCTGGACATGT	996
Db	1095	CGCCAAGACATTTCCAAATGGTGTGTTTCCATAAAGGGCTCCGCAATCGCTGGACCTGC	1154
Qy	997	ATTGGAGACCTAATCGGAGTCCACACCAAGCCTTCAGAAGTGGAGGATTCATTTGTACC	1056
Db	1155	ATTGGAGACCTAATCGAAGCCTTACACCAAGCCTTAAGAGGTGGAGGATTCATCTGTACA	1214
Qy	1057	CAGAAATGGCAATTTTACCAAGCATTTTCAGGATTTAGTATTATCTATGAAAGCTGTAG	1116
Db	1215	AAGGATCACTACATTTTACAGGCATTTCAAAATATATCTCCGTTATGGGTTCTGTAG	1274
Qy	1117	TAAACTTGGTGAAGGACACAGGTACTATCATTTGAAAACTTGACAAATGGGTTCTTCTCC	1176
Db	1275	TAAACTCGGTGAAGGCCACACCCCTCTGCTCTTGAACACACTGGCACTGGAACATCTCGC	1334
Qy	1177	ATTACACC--TTCTTTATATTTTAAAGGCGCTGTGAAT	1211
Db	1335	CTTGGATCTGTCTTCTCATAATTTTCAAGGCTTCTGAGT	1371

157	QY	TGGTTTACTCTTTTATAAGTCTACCTTAAAGACAAACACAGGAAGTGGAGAGACTGGTTA	216
324	Db	TGGTTTATCTTTTATAAGTTACCCAAAAGGACTAGCAAGGCACAGTGAAGAGCGGGGGTG	383
217	QY	GAGTACCTGTACCTAGAGCTCTACAACCTAGAAGCTGGAGGAGAGTGGACACTAATGAAT	276
384	Db	CAGTACCTGTACTTGGACTCCACAAGACAAACCTTGGACAGAGCCTCTACCTGATTAAC	443
277	QY	GACACCAAGAGTGTTTTGGGAAGGACATTACAACAGCTATATGAAGCATATGCCCTCTAAG	336
444	Db	AGCACCAAGAGTGTCTTGGGAGGACCTTACAGCATCTGTATGCACACATAATTTCCAG	503
337	QY	AGTAACAACACAGCCTCTCTAATATACAATGATGGAGTCCCTAAACCTGTGAATTACAGT	396
504	Db	AATGACACACAGCCTCTCTAATATACAACGATGGTGTCCCTGGATCTGTGAATTTACAGC	560

```

; LENGTH: 1224
; TYPE: DNA
; ORGANISM: Mus sp.
US-09-574-942-1

Query Match      53.1%; Score 672.8; DB 4; Length 1224;
Best Local Similarity 75.0%; Pred. No. 2.1e-206;
Matches 871; Conservative 0; Mismatches 282; Indels 9; Gaps 2;

QY 37 AAACAGAAATGATGGCAAGCTCTGTAAGACATCTTTGCTTTCCTTGGGCTC 96
Db 37 AAACAGAAATGATGGCAAGCTCTGTAAGACATCTTTGCTTTCCTTGGGCTC 96
QY 31 AGATAGAAATGACAGCAAGCTCTAAGACAGTCTTTCTTCTTTCCTTGGGCTC 90
Db 31 AGATAGAAATGACAGCAAGCTCTAAGACAGTCTTTCTTCTTTCCTTGGGCTC 90
QY 97 TTTGGGGTGGTGGGGCAGCAACATTTTCATGCAAGAAATGAAGAGGAAAGCTGTGGAC 156
Db 97 TTTGGGGTGGTGGGGCAGCAACATTTTCATGCAAGAAATGAAGAGGAAAGCTGTGGAC 156
QY 91 TCTGGGGTCTGGGGCAGCAACAAATCTCATGCAAGAAATGAATATGTTGAAGCTGTGGAC 150
Db 91 TCTGGGGTCTGGGGCAGCAACAAATCTCATGCAAGAAATGAATATGTTGAAGCTGTGGAC 150
QY 157 TGGTTTACTTTTTTATAAGTTTACCTTAAAGACAAACCAAGAAAGTGGAGACTGGGTTA 216
Db 157 TGGTTTACTTTTTTATAAGTTTACCTTAAAGACAAACCAAGAAAGTGGAGACTGGGTTA 216
QY 151 TGGTTTACTTTTTTATAAGTTTACCTTAAAGACAAACCAAGAAAGTGGAGACTGGGTTG 210
Db 151 TGGTTTACTTTTTTATAAGTTTACCTTAAAGACAAACCAAGAAAGTGGAGACTGGGTTG 210
QY 217 GAGTACCTGTACCTAGACTCTAACAAGTGAAGCTGGAGAGAGTGAACACTTAATGAAT 276
Db 217 GAGTACCTGTACCTAGACTCTAACAAGTGAAGCTGGAGAGAGTGAACACTTAATGAAT 276
QY 211 CAGTACCTGTACCTGGACTTCCACAGCAAAACCTTGGACAAAGAGCTCTACCTGATTAA 270
Db 211 CAGTACCTGTACCTGGACTTCCACAGCAAAACCTTGGACAAAGAGCTCTACCTGATTAA 270
QY 277 GACACCAAGAGTGTTTGGGAAGGACATTACAAAGCTATATGAAGCATATGCCTCTAAG 336
Db 277 GACACCAAGAGTGTTTGGGAAGGACATTACAAAGCTATATGAAGCATATGCCTCTAAG 336
QY 271 AGCACCAGAGGTGCTCTGGGAGAGACCTTACAGCATCTGATGACACATATTCACG 330
Db 271 AGCACCAGAGGTGCTCTGGGAGAGACCTTACAGCATCTGATGACACATATTCACG 330
QY 337 AGTAACACACAGAGCTATCTAATATACAAATGATGGAGTCCCTTAAACCTGTGAATPACGT 396
Db 337 AGTAACACACAGAGCTATCTAATATACAAATGATGGAGTCCCTTAAACCTGTGAATPACGT 396
QY 331 AATG---ACACAGCTATCTAATATACAAAGATGTGTCCTGGATCTGTGAATTTACAGC 387
Db 331 AATG---ACACAGCTATCTAATATACAAAGATGTGTCCTGGATCTGTGAATTTACAGC 387
QY 397 AGAAGTATGACACACCAAGAGTTTACTGTGTGGACAGAGTTCAAGGTTCTGGCTG 456
Db 397 AGAAGTATGACACACCAAGAGTTTACTGTGTGGACAGAGTTCAAGGTTCTGGCTG 456
QY 388 AGACAGTATGACATGCCAAAGGTCTGCTGTATGGAACAGAAACAGAGGTTCTGGCTG 447
Db 388 AGACAGTATGACATGCCAAAGGTCTGCTGTATGGAACAGAAACAGAGGTTCTGGCTG 447
QY 457 ATTCAATTCATCTCTCAGTTTCTTCAATTCGGGAAGAGGCTATGATTATCCACCCACA 516
Db 457 ATTCAATTCATCTCTCAGTTTCTTCAATTCGGGAAGAGGCTATGATTATCCACCCACA 516
QY 448 ATACACTCTGTCCCAAGTTTCCCCAGTTC-----ATGGCTATGAGTACCCAACTCG 501
Db 448 ATACACTCTGTCCCAAGTTTCCCCAGTTC-----ATGGCTATGAGTACCCAACTCG 501
QY 517 GGGAGACAAATGGAACAAAGTGGCATCTGCATAAATTTTCAAGTACAAACAGATGAGGCA 576
Db 517 GGGAGACAAATGGAACAAAGTGGCATCTGCATAAATTTTCAAGTACAAACAGATGAGGCA 576
QY 502 GGGAGGCATATGACAAACCGCATCTGCATCACTTTCCGATACAGACAGTTCGAGAA 561
Db 502 GGGAGGCATATGACAAACCGCATCTGCATCACTTTCCGATACAGACAGTTCGAGAA 561
QY 577 ATAGATTCAGCTCTGCTGCAACCCACAGTCTATAGCTGCTCCATCCACCCACC 636
Db 577 ATAGATTCAGCTCTGCTGCAACCCACAGTCTATAGCTGCTCCATCCACCCACC 636
QY 562 ATAGATTTTCACTCTTGGTCTTACAAACCAACATCTACAGCTCTTCAATCCAAAGCACC 621
Db 562 ATAGATTTTCACTCTTGGTCTTACAAACCAACATCTACAGCTCTTCAATCCAAAGCACC 621
QY 637 TTTTACACAGGAGCTCATTCATATGCCAGCTGTGCACAGGGCCAGCTCATCAGAGATT 696
Db 637 TTTTACACAGGAGCTCATTCATATGCCAGCTGTGCACAGGGCCAGCTCATCAGAGATT 696
QY 622 TTTTACTGGAACCTTATCTATGATCCCGGATGTGTGCCACTCCAGTTCTTAAAGATC 681
Db 622 TTTTACTGGAACCTTATCTATGATCCCGGATGTGTGCCACTCCAGTTCTTAAAGATC 681
QY 697 CTGCGAGGCTCTTCCACCACTCTCAGTGGGCCCAGGACAAAATTCCTCATTTTGA 756
Db 697 CTGCGAGGCTCTTCCACCACTCTCAGTGGGCCCAGGACAAAATTCCTCATTTTGA 756
QY 682 CTTGTCGGTACCTCGCTGAACTGCACTCAGCCAGGGTCTAAACTTGTCTCATTTTGA 741
Db 682 CTTGTCGGTACCTCGCTGAACTGCACTCAGCCAGGGTCTAAACTTGTCTCATTTTGA 741
QY 757 AAGTCGGATCTTTTCTGACACATCTTTCAGCTGATGCTCAAGCTCAAGCTGAAGACA 816
Db 757 AAGTCGGATCTTTTCTGACACATCTTTCAGCTGATGCTCAAGCTCAAGCTGAAGACA 816
QY 742 AAATCAAGTTTTTATCTACTGATGACATCTTTACAGATGATAGTCAAAAGTTGAAGACA 801
Db 742 AAATCAAGTTTTTATCTACTGATGACATCTTTACAGATGATAGTCAAAAGTTGAAGACA 801
QY 817 CACTTGTGTAAACAGAAACCTGGCAGGAAAGACAAAGACTTCTTCAAACTGCTCCCTT 876
Db 817 CACTTGTGTAAACAGAAACCTGGCAGGAAAGACAAAGACTTCTTCAAACTGCTCCCTT 876
QY 802 CATTGTGTAGCACAACTGGCAGAAAGAAACAAAGAGTCTTCTTCAAACTGTTCCCTG 861
Db 802 CATTGTGTAGCACAACTGGCAGAAAGAAACAAAGAGTCTTCTTCAAACTGTTCCCTG 861
QY 877 CTTTACCAATGTCTAATAATAAAGCAATTAATATACGACACTCTTATTTAGTTCT 936
Db 877 CTTTACCAATGTCTAATAATAAAGCAATTAATATACGACACTCTTATTTAGTTCT 936
QY 862 CTTTACCAATGTCTAATAATAAAGCAATTAATTTGGGGTAACTTCAAGTCTTACTTCAGTTCT 921
Db 862 CTTTACCAATGTCTAATAATAAAGCAATTAATTTGGGGTAACTTCAAGTCTTACTTCAGTTCT 921
QY 937 TATCAGATCAGCCAGGTGTATTTCCAAAAGGGCACCACAAATCGCTGGACATGT 996
Db 937 TATCAGATCAGCCAGGTGTATTTCCAAAAGGGCACCACAAATCGCTGGACATGT 996
QY 922 CGCCAGACCAATTCCAAATGGGTGTCTTCCATAAAGGGCTCCGCAATTCGCTGGACCTGC 981
Db 922 CGCCAGACCAATTCCAAATGGGTGTCTTCCATAAAGGGCTCCGCAATTCGCTGGACCTGC 981

```

```

RESULT 5
US-09-147-915-1
; Sequence 1, Application US/09147915A
; Patent No. 6184034
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan
; APPLICANT: Krissner, Ronald
; TITLE OF INVENTION: Deoxyribonuclease II Proteins and cDNAs
; FILE REFERENCE: DC-0097
; CURRENT APPLICATION NUMBER: US/09/147,915A
; CURRENT FILING DATE: 1999-03-23
; EARLIER APPLICATION NUMBER: PCT/US97/18262
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/028,539
; EARLIER FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1915
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-147-915-1

```

```

Query Match      6.6%; Score 84; DB 3; Length 1915;
Best Local Similarity 47.5%; Pred. No. 1.8e-16;
Matches 471; Conservative 0; Mismatches 485; Indels 36; Gaps 6;

QY 142 GGGAAAGCTGTGGACTGGTTTACTTTTATAAGTT---ACCTAAAGACAAAACAAGGAA 198
Db 137 GGGAGCCTGTAGACTGGTTTCTGTCTACAGCTGCCAGCTCTTAGAGGGTCCGGGGAG 196
QY 199 AGTGAGAGACTGGTTAGAGTACCTGTACCTAGACTCTACAACTAGAGCTGGAGGAG 258
Db 197 GCGGCGCAGAGAGGGCTGCAGTACAAAGTATCTGGACAGAGCTCCGGAGGCTGGCGGGAC 256
QY 259 AGTGAGCACTAATGAANTGACACCAAGAGTGTTTTGGGAAGACATTACACAGCTATAT 318
Db 257 GCGAGGCACTCATCACAAGCCGGAGGGGCGCGTGGCCGAGGCTGCAGCGCTGTAC 316
QY 319 GAAAGCATATGCTCTPAGAGTAACAAACAGCCTTATCTAATATATAAATGATGAGTCCCT 378
Db 317 CGG-----AGCAACACAGCCAGCTGCGCTTCTCTCTCTCAATGATGACCAACCGCT 367
QY 379 AAACC-----TGTGAATACAGTAGAAGATGACACACAAAGGTTTACTGCTGTGG 432
Db 368 CAAACAGCAAGGCTCAGGACTCTTCCATGCGTGGGACACAAAGGTTGCTGCTCTCTT 427
QY 433 AACAGAGTTCAAGGGTCTTGGCTGATTTCATTCATCCCTCAGTTTCTTCCCAATTCGGAA 492
Db 428 GACACAGTGGGGCTTCTGGCTGTCCACAGTGTACCTACTTCCCTCCACCGGCTCC 487
QY 493 ---GAAGGCTATGATTATCCACCCACAGGAGACGAATGGAACAAGTGGGCTCTGCATA 549
Db 488 TCTGCTGCATACAGTTGGCTCTATAGCGCTCTAGCTAGCGGACAGCCCTGCTCTGTGTG 547
QY 550 ACTTTCAAGTACCAACCAAGTATGAGCAATAGATTCTAGCTCTTGTGCTGCAACCCCAAC 609

```

Fri Oct 15 10:01:05 2004

```

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1575 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: DNA
US-08-639-294-1

Query Match      6.5%; Score 82.4; DB 3; Length 1575;
Best Local Similarity 47.4%; Pred. No. 5.2e-16;
Matches 470; Conservative 0; Mismatches 486; Indels 36; Gaps 6;

142 GCGAAGAGCTGTGACCTGGTTTACTTTTATPAAGTT---ACCTAAAGACAAAACAGGAA 198
155 GGGCAGCCTGTAGACTGGTTTCGTGGTCTACAAGTCCAGCTCTTAGAGGGTCCGGGGAG 214
199 AGTGGAGAGACTGGGTTAGAGTACCTGTACTAGACTCTACAAGTGAAGCTGGAGGAAG 258
215 GCGGCGCAGAGAGGGCTGCAGTATCTGGACGAGAGCTCCGAGAGCTGGCGGGAC 274
259 AGTCAGCAACTAATGAATGACACCAAGAGTGTGTTGGGAAGGACATTTACACAGCTATAT 318
275 GGCAGGGCACTTCATCAACAGCCCGAGGGGGCCCTGGCCGGAAGCTGCAGCCGTGTAC 334
319 GAAGCATATGCTCTAAGAGTAAACAACAGCCCTATCTAATATACATGATGGAGTCCCT 378
335 CGG-----AGCAACACACAGCCAGCTCGCTTCTCTTCTTCAATGACCAACCGCT 385
379 AAAC-----TGTAATTCAGTAGAAGTATGACACACCAAGGTTTACTGCTGTGG 432
386 CAACCCAGCAAGGCTCAGGACTCTTCCATGCGTGGGACACGAAGGGTGTCTCTCTCT 445
433 AACAGAGTTCAAGGTTCTGCTGATTCATTCCTCCTCAGTTTCCCTCAATTCGGAA 492
446 GACCAAGATGGGGCTTCTGCTGGTCCACAGTGTACCTTAACCTTCCCTCCACCGGCTCC 505
493 ---GAAGGCTATGATTAATCCACCCAGGAGAGCAATGGAAGGAGGATCTGCATA 549
506 TCTGCTGCATACAGCTGGCTCATAGCGCTGTACCTACGGGCGAGACCTGCTCTGTGTG 565
550 ACTTTCAGAGTACACAGTATGAGGCAATAGATTCTCAGCTCTTGTGCAACCCCAAC 609
566 TCTTTTCCCTTTCGCTCAGTTCTCGAAGATGGGCAAGCAGCTACCTACACCTCCCTGG 625
610 GTCTATAGTCTCCATCCAGCCACCTTTTACACAGGAGCTCAATTCACATGCCCGAGTG 669
626 GTCTATTAACCTACAGCTGGAAGGATCTTTGCCAGGAATTCCTCCGACTTGGAGAGTGTG 685
670 TGCACCCAGGGCCAGCTCATFAGAGATTCCTGGCAGGCTCTTCCACACACTTCAGTGGCC 729
686 GTCAAGGGCCACACGTTAGCCCAAGAACCTTGGAAACAGCAGCATCACATCCCAG 745
730 CAGGGACAAAATTCCTCCATTTTCAAGTCGGATTCTTTTCTTGACGACATCTTTGCA 789
746 GCGGGGGCTGTTTTCAGAGCTTTGCCAAGTTTCAGCAAAATTTGGAGATGACCTGTACTCC 805
790 GCTGGATGCTCAACGGCTGAAGACACACTTGTGTTAACAGAAACCTGGCAGCGAAAAAGA 849
806 GGCTGTTGGCAGCAGCCCTTGGTACCACTTCAGCTCCAGTTCAGTTCGGCAGCAAACTGTA 865
850 CAAGAGCTTCTTCAAACTGCTCCCTTTCCTTACCATGTCTACATATATAAGCAATTA 909
866 GGCACTCTGCTTAACTGCTCGGATATCTGGCAGGTTCTGAATGTGAACCAAGTAGCT 925
910 TTATACGACA-----CTCTTATTTTCAAGTCTTATCAAGATCACCGCAGTGGTGTATT 963
926 TTCCCTGGACCCAGCGGCCCAAGCTTCAACAGCAGAGGACCACTTCCAAATGTTGGGTG 985
964 TCCCAAAAGGGCCACCAAAAATCGCTGACATGTATTGGAGACCTTAATTCGAGTCCACAC 1023
986 TCCCAAAAGGGCC-----CTGGACCTGCGTGGTACATGATCGGAACCCAGGGA 1036

```

RESULT 6

```

US-08-639-294-1
; Sequence 1, Application US/08639294
; Patent No. 6265195
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Baron, Will F.
; TITLE OF INVENTION: Human DNase II
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/639,294
; FILING DATE: 25-Apr-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Johnston, Sean A.
; REGISTRATION NUMBER: 35,910
; REFERENCE/DOCKET NUMBER: P1024
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3562
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168

```

QY 1024 CAAGCCTTCAGAGTGGAGGATTCATTTGTATCCAGAAATGGCAATTTACCAAGCAATTT 1083  
 Db 1037 GAGGAGCAACGGGTGGGGGACACACTGTGTGCCAGCTGCCAGCCCTCTGGAAGCCCTTC 1096  
 QY 1084 CAAGGATTAGTATTATATATATGAAAGCTGTAA 1115  
 Db 1097 CAGCCGCTGGTGAAGAACTACCAAGCCCTGTAA 1128

## RESULT 7

US-09-861-034B-1  
 ; Sequence 1, Application US/09861034B  
 ; Patent No. 6569429

## GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
 ; Baron, Will F.  
 ; TITLE OF INVENTION: Human DNase II  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Genentech, Inc.  
 ; STREET: 1 DNA Way  
 ; CITY: South San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94080

## COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: WinPatIn (Genentech)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/861,034B  
 ; FILING DATE: 18-May-2001  
 ; CLASSIFICATION: <Unknown>

## PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/639294  
 ; FILING DATE: 25-APR-1996

## ATTORNEY/AGENT INFORMATION:

; NAME: Johnston, Sean A.  
 ; REGISTRATION NUMBER: 35,910  
 ; REFERENCE/DOCKET NUMBER: P1024D1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 650/225-3562  
 ; TELEFAX: 650/952-9881

## INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1575 base pairs  
 ; TYPE: Nucleic Acid  
 ; STRANDEDNESS: Single  
 ; TOPOLOGY: Linear  
 ; MOLECULE TYPE: DNA

; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-861-034B-1

Query Match 6.5%; Score 82.4; DB 4; Length 1575;  
 Best Local Similarity 47.4%; Pred No. 5.2e-16;  
 Matches 470; Conservative 0; Mismatches 486; Indels 36; Gaps 6;

QY 142 GGGAAAGCTGTGGCTGTGTTTACTTTTATAAGTT---ACCTAAAGACAAACAAAGGAA 198  
 Db 155 GGGCAGCCTGTAGACTGGTCTGTTCTACAAAGTCCAGCTCTTAGAGGTCGGGGAG 214  
 QY 199 AGTGGAGAGACTGGTTAGTAGTACCTGTACTAGACTCTCAACTAGAACTGAGGAAG 258  
 Db 215 GCGCGCAGAGAGCGCTGCGAGTACAGTATCTGACGAGAGCTCCGAGGCTGGCGGAC 274  
 QY 259 AGTGAGCAACTAATGAATGACCAAGAGTGTTTGGGAAGGACATTACACAGCTATAT 318  
 Db 275 GGCAGGCACTCATCAACAGCCGGGGGGCGGTGGCCGAGAGCTGACGCGCTGTAC 334  
 QY 319 GAACCATATGCTTAAAGATTAACACAGAGCTATCTTAATATACATGATGAGTCCCT 378  
 Db 335 CGG-----AGCAACACAGCCAGCTCGCCTTCTGCTCTTACATGACCAACCGCT 385

QY 379 AAAC-----TGTGAATTACAGTAGAAAGTATGACACACAAAGTTTACTGTGTGG 432  
 Db 386 CAACCCAGCAAGGCTCAGGACTCTTCATGCTGGGCACACGAGGGTGTCTGCTCTTT 445  
 QY 433 AACAGAGTTCAGGGTTCCTGCTGATTCATTCCTCAGTTCCTTCCATTTCCAAATTCGGAA 492  
 Db 446 GACCAAGATGGGGCTTCTGCTGGTCCACAGTGTACTTACTTCCCTCCACCGGCTCC 505  
 QY 493 ---GAAGGCTATGATTATCAACCCACAGGAGACGAATGGACAAAGTGGCATCTGCATA 549  
 Db 506 TCTGCTGCATACAGCTGGCTCTAGCGCTGTACCTACGGGCAGACCCCTGCTCTGTGTG 565  
 QY 550 ACTTTCAGTACACAGTAGTAGGCAATAGATTCTCAGCTCTTTGGTCTGCAACCCCAAC 609  
 Db 566 TCTTTTCCCTTCGCTCAGTTCTCGAAGATGGGCAAGCAGCTGACCTACACCTACCCCTGG 625  
 QY 610 GTCTATAGCTGCTCCATCCAGCCACTTTTCCACGAGAGCTCATTCACATGCCCCAGCTG 669  
 Db 626 GTCTATACCTACAGCTGGAGGGATCTTTGCCAGGAATTTCCCGACTTGGAGATGTG 685  
 QY 670 TGCACAGGCGCCAGCTCATCAGAGATTCTTGGCAGGCTCTCTCACACACTTCAGTCCGCC 729  
 Db 686 GTCAAGGGCCACCACTGTAGCCAGAACCTCTGGAACAGCAGCATCACTCAATCCAG 745  
 QY 730 CAGGACAAATAATCTCTCATTTTGCAGAGTGGATTTCTTTCTTGACGACATCTTTGCA 789  
 Db 746 GCGGGGCTGTTTTCCAGAGCTTTGCCAAGTTTGCAGAAATTTGGAGATGACCTGTACTCC 805  
 QY 790 GCTGTGATGGCTCAACGGCTGAAGACACACTTGTAAACAGAAACCTGGCAGCGCAAAAGA 849  
 Db 806 GGTGTGGCAGAGCCCTTGTGTACCACTGAGGTCAGTTCTGGCACAACCTGTA 865  
 QY 850 CAAGAGCTTCTTCAAACTGCTCCCTTCTTACCATGTCTCAATATAAAAGCAATTA 909  
 Db 866 GGCATCTTGCCTCTTAACCTGCTCGGATATCTGGCAGTTCTGATGTGAACAGATAGCT 925  
 QY 910 TTATCAGACA-----CTCTATTTCAGTTCTTATCAAGATCAAGCAGTGGTGTATT 963  
 Db 926 TTCCCTGGACACAGCGCGCCCAAGCTTCAACAGCAGAGGACCACTTCCAAATTTGGTGG 985  
 QY 964 TCCCAAAAGGCGCAACAAATCGCTGGACATGTATTGGAGACTTAATCGGAGTCCACAC 1023  
 Db 986 TCCCAAAAGGCGC-----CTGGACCTGCGTGGTGACATGATCGGAACCAAGGA 1036  
 QY 1024 CAAGCTTTCAGAGTGGAGGATTCATTTGTATCCAGAAATTTGCAATTTTCCACCAATTT 1083  
 Db 1037 GAGGAGCAACGGGTGGGGGCACACTGTGTGCCAGCTGCTGCCAGCTGCCAGCCCTCTG 1096  
 QY 1084 CAAGGATTAGTATTATATATGAAAGCTGTAA 1115  
 Db 1097 CAGCCGCTGGTGAAGAACTACCAAGCCCTGTAA 1128

## RESULT 8

US-09-147-915-2

; Sequence 2, Application US/09147915A

; Patent No. 6184034

## GENERAL INFORMATION:

; APPLICANT: Eastman, Alan  
 ; APPLICANT: Krieser, Ronald  
 ; TITLE OF INVENTION: Deoxyribonuclease II Proteins and cDNAs  
 ; FILE REFERENCE: DC-0097  
 ; CURRENT APPLICATION NUMBER: US/09/147,915A  
 ; CURRENT FILING DATE: 1999-03-23  
 ; EARLIER APPLICATION NUMBER: PCT/US97/18262  
 ; EARLIER FILING DATE: 1997-10-09  
 ; EARLIER APPLICATION NUMBER: 60/028,539  
 ; EARLIER FILING DATE: 1996-10-15  
 ; NUMBER OF SEQ ID NOS: 18  
 ; SOFTWARE: Patent in Ver. 2.0  
 ; SEQ ID NO 2  
 ; LENGTH: 927



```

; TYPE: DNA
; ORGANISM: Bos sp.
US-09-147-915-2

Query Match
Best Local Similarity 5.6%; Score 71; DB 3; Length 927;
Matches 317; Conservative 0; Mismatches 330; Indels 18; Gaps 3;

QY 405 TGGACACACAAAGTTTACTGCTGTGGACACAGATTCAAGGTTCTGGCTGATTCTTC 464
Db 75 TGGGACACAGAAAGGTGTGCTCTCTGGACCAAGAGGGGGTCTTGGTTGATCCACAG 134
QY 465 CATCCCTCAGTTTCTCCCAATTCGGG---AGAAGGCTATGATTATCCACCCACAGGGAG 521
Db 135 CGTTCCAAACTTCCCTCCACGTCCTCTCTGCTGGTACAGCTGGCCCTCTGGTGCCCA 194
QY 522 ACGAATGACAAAGTGGCATCTGCATAAATTTCAAGTACAAACAGTATGAGCAATAGA 581
Db 195 AAAATATGGGACACCTGATCTGTGTAATTTTCTCTCACCCAGTTCTCTGGATATCAG 254
QY 582 TTCTCAGCTCTTGGTCTGCAACCCCAAGTCTATAGTCTCATCCAGCCACCTTTCA 641
Db 255 CAACAGCTGACCTATACCTATCCACTGGTATATGACACAGGCTGGAAGGGGACTTTGG 314
QY 642 CCAGGAGCTCATTCACATGCCAGCTGTGCACCAGGCCAGCTCATCAGAGATTCTCTGG 701
Db 315 CAGAAATTCCTACTCTGGAGAGGTAGTCAAGGGCCATCAGTTGCCAGGAGCGTG 374
QY 702 CAGGCTCTCACACACTTCAGTCCGCCAGGACCAAAATTTCTCATTTTCCAAAGTC 761
Db 375 GAACAGCAGTGTAACTCATCAATCAAGAAAGGAGCCACATTCAGAGCTTTGCCAAAT 434
QY 762 GGATCTTTTCTTGACAGACATCTTTGAGCCTGGATGGCTCAACGGTCAAGACACATT 821
Db 435 TGAACATTTTGGAGATGACCTGTACTCTGGCTGGCTGGCGAAGCCCTTTGGCAGTACCCT 494
QY 822 GTTAAACAGAACTGGCAGCGAAAGACAGAGCTTCTTCAAACTGTCTCCCTTCTCTTA 881
Db 495 GCAGGTCCTAATCTGGCAAGATCTTCTGGTATCTTGGCTTCACTGCTCTGGGGCCCA 554
QY 882 CCAATGCTCAATATAAAGCAATTAATATCAAGACACTCT-----TATTTCAGTTC 935
Db 555 GCATGATATTGAGTGTACTCAGACAGCTTTCCCTGGCCAGCTGGGCCAGCTTCAATGC 614
QY 936 TTATCAGATCAGCCCAAGTGGTGTATTTCCCAAGGGGACCAAAATCGCTGGACATG 995
Db 615 CACAGAAGACCAATTCCAAGTGGTGTATACCCCAAGGGGCC-----CTGGGCGTG 665
QY 996 TATTGGAGACCTAAATCGGAGTCCACACCAAGCTTCCAGAAAGTGGAGATTCTATTGTAC 1055
Db 666 TGTGGTGAATGAATCGGAACCAAGAGAGGAGCACCGGGGTGGGGGCACTGTGTGTC 725
QY 1056 CCAGA 1060
Db 726 CCAGA 730

```

```

RESULT 9
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA

```

```

; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)883-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
; US-08-232-463-14

Query Match 4.1%; Score 51.8; DB 1; Length 7218;
Best Local Similarity 7.9%; Pred. No. 1e-05;
Matches 35; Conservative 217; Mismatches 189; Indels 0; Gaps 0;

QY 100 GGGGTCCTGGGGCAGCAACAATTCATGAGAAATGAAGAGGAAAGCTGTGGACTGG 159
Db 1423 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1364
QY 160 TTTACTTTTTTAAAGTTACCTAAAGACAAACAAAGAGTGGAGAGACTGGTTAGAG 219
Db 1363 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1304
QY 220 TACTGTACTAGACTCTACACTAGAGTGGAGAGAGTGGAGCAACTAATGAATGAC 279
Db 1303 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1244
QY 280 ACCAAGAGTGTTTTGGGAAGGACATTACACAGCTATATGAAGCATATGCTCTAAGAGT 339
Db 1243 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1184
QY 340 AACACACAGCTTATCTAATATACATGATGGAGTCCCTAAACCTGTGAATTACAGTAGA 399
Db 1183 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1124
QY 400 AAGTATGGACACACAAAGGTTTACTGCTGTGGAACAGAGTTCAAGGGTTCTGGCTGATT 459
Db 1123 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1064
QY 460 CATTCATCCCTCAGTTTCTCCAAATTCGGAAGAAGGCTATGATTATCCACCACAGGG 519
Db 1063 GCAAGTCCCTCGACCTGGAGCCAGCTCGAATTAATCTGTGAGGCTATGCAACGA 1004
QY 520 AGACGAATAGCAAAAGTGGC 540
Db 1003 AGGAAATAGTTATAGTAGC 983

```

```

RESULT 10
US-09-621-976-15639/c
; Sequence 15639, Application US/09621976

```



NAME/KEY: intron  
LOCATION: (136262)...(137936)  
NAME/KEY: exon  
LOCATION: (137937)...(138053)  
NAME/KEY: intron  
LOCATION: (138054)...(138637)  
NAME/KEY: exon  
LOCATION: (138638)...(138766)  
NAME/KEY: intron  
LOCATION: (138767)...(138864)  
NAME/KEY: exon  
LOCATION: (138865)...(138940)  
NAME/KEY: intron  
LOCATION: (138941)...(139765)  
NAME/KEY: exon  
LOCATION: (139766)...(139860)  
NAME/KEY: intron  
LOCATION: (139861)...(142245)  
NAME/KEY: exon  
LOCATION: (142246)...(142445)  
NAME/KEY: intron  
LOCATION: (142446)...(143605)  
NAME/KEY: exon  
LOCATION: (143606)...(143738)  
NAME/KEY: intron  
LOCATION: (143739)...(145838)  
NAME/KEY: exon  
LOCATION: (145839)...(145931)  
NAME/KEY: intron  
LOCATION: (145932)...(147385)  
NAME/KEY: exon  
LOCATION: (147386)...(147544)  
NAME/KEY: intron  
LOCATION: (147545)...(153274)  
NAME/KEY: exon  
LOCATION: (153275)...(153321)  
NAME/KEY: intron  
LOCATION: (153322)...(155088)  
NAME/KEY: exon  
LOCATION: (155089)...(155231)  
NAME/KEY: intron  
LOCATION: (155232)...(156025)  
NAME/KEY: exon  
LOCATION: (156026)...(156151)  
NAME/KEY: intron  
LOCATION: (156152)...(156826)  
NAME/KEY: exon  
LOCATION: (156827)...(156928)  
NAME/KEY: intron  
LOCATION: (156929)...(163399)  
NAME/KEY: exon  
LOCATION: (163400)...(163586)  
US-09-676-610B-24

Query Match 3.1%; Score 39; DB 4; Length 169998;

Best Local Similarity 49.3%; Pred. No. 1.1; Mismatches 105; Indels 0; Gaps 0;

Matches 102; Conservative 0

QY 1052 GTACCCAGAAATGGCAAAATTTACCAAGCAATTTCAAGGATTAGTATTATATCTATGAAAGCT 1111  
DB 56948 GAACACCGTCTGGTAATTAAGATGACAGGTTAAGCAATTTCTTCACTACCATAAGG 56889  
QY 1112 GTAAGTAAACTTGGTGAAGACACAGGTACTATCATTTGAAACCTTGACAATGGGTCTT 1171  
DB 56888 AAAAATAATCTTGCTTTGGCCACTAATTTGATTGCTTACTTAATTTGATTTTAAATGTCTT 56829  
QY 1172 CTTCATTACACCTCTTTTATTTTAAAGCCGTGTAATATCTTATATACCTTGCAATC 1231  
DB 56828 TTCACCTTAAGAGTGTCTTATTTACAGTGGTAGTTTATAAGGTTAAAGTATGATCTCC 56769  
QY 1232 ACAAAATAAAACATATTTCTCTCATGT 1258  
DB 56768 CAAATATTAGATATCTTTAATCATGT 56742

## RESULT 13

US-09-877-177A-10/c

Sequence 10, Application US/09877177A  
Patent No. 6582919  
GENERAL INFORMATION:  
APPLICANT: K. Danenberg  
TITLE OF INVENTION: Method of determining Epidermal Growth  
TITLE OF INVENTION: Factor Receptor and HER2-Neu Gene Expression  
TITLE OF INVENTION: and Correlation of Levels Thereof With Survival  
FILE REFERENCE: 11220/120  
CURRENT APPLICATION NUMBER: US/09/877,177A  
CURRENT FILING DATE: 2001-06-11  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10  
LENGTH: 197496  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-877-177A-10

Query Match 3.1%; Score 39; DB 4; Length 197496;

Best Local Similarity 49.3%; Pred. No. 1.2;

Matches 102; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 1052 GTACCCAGAAATGGCAAAATTTACCAAGCAATTTCAAGGATTAGTATTATATCTATGAAAGCT 1111

DB 64948 GAACACCGTCTGGTAATTAAGATGACAGGTTAAGCAATTTCTTCACTACCATAAGG 64889

QY 1112 GTAAGTAAACTTGGTGAAGACACAGGTACTATCATTTGAAACCTTGACAATGGGTCTT 1171

DB 64888 AAAAATAATCTTGCTTTGGCCACTAATTTGATTGCTTACTTAATTTGATTTTAAATGTCTT 64829

QY 1172 CTTCATTACACCTCTTTTATTTTAAAGCCGTGTAATATCTTATATACCTTGCAATC 1231

DB 64828 TTCACCTTAAGAGTGTCTTATTTACAGTGGTAGTTTATAAGGTTAAAGTATGATCTCC 64769

QY 1232 ACAAAATAAAACATATTTCTCTCATGT 1258

DB 64768 CAAATATTAGATATCTTTAATCATGT 64742

## RESULT 14

US-09-134-000C-3209/c

Sequence 3209, Application US/09134000C

Patent No. 6617156

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 032796-032

CURRENT APPLICATION NUMBER: US/09/134,000C

CURRENT FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/055,778

PRIOR FILING DATE: 1997-08-15

NUMBER OF SEQ ID NOS: 6812

SOFTWARE: PatentIn version 3.1

SEQ ID NO 3209

LENGTH: 1509

TYPE: DNA

ORGANISM: Enterococcus faecalis

US-09-134-000C-3209

Query Match

Best Local Similarity 2.7%; Score 34.6; DB 4; Length 1509;

Matches 73; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 1103 ATGAAGCTGTAGTAAACTTGGTGAAGACACAGGTACTATCATTTGAAACCTTGACA 1162

DB 1110 ATTTAAACCTTGAATTAACCTCTGCTAAAGAACCGGATATTTTCTTTGATTTTATTTCCA 1051

QY 1163 ATGGGTCTTCTCCATTACACCTTCTTTATATTTTAAAGCCCTGTGAATATCTTATAAC 1222

Db 1050 AAAGACATTGACATTTTCACATTTTGACATAGTAAAGTTTCGGCAATAAATGTGTAGC 991  
 QY 1223 CTGGATATCACAATA 1239  
 Db 990 AAACAAATAATAAAAA 974

RESULT 15

US-08-484-438-5  
 ; Sequence 5, Application US/08484438  
 ; Patent No. 5811098  
 ; Patent No. 5811098 5780031  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Plowman, Gregory D.  
 ; APPLICANT: Culouscou, Jean-Michel  
 ; APPLICANT: Shoyab, Mohammed  
 ; APPLICANT: Siegall, Clay B.  
 ; APPLICANT: Hellistr m, Ingegerd  
 ; APPLICANT: Hellistr m, Karl E.  
 ; TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE  
 ; NUMBER OF SEQUENCES: 42  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10036-2711  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/484,438  
 ; FILING DATE: 07-JUN-1995  
 ; CLASSIFICATION: 530  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/323,442  
 ; FILING DATE: 14-OCT-1994  
 ; APPLICATION NUMBER: US 08/150,704  
 ; FILING DATE: 10-NOV-1993  
 ; CLASSIFICATION: 530  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/981,165  
 ; FILING DATE: 24-NOV-1992  
 ; CLASSIFICATION: 530  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Mistrock, S. Leslie  
 ; REGISTRATION NUMBER: 18,872  
 ; REFERENCE/DOCKET NUMBER: 5624-230  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 790-9090  
 ; TELEFAX: (212) 869-8864/9741  
 ; TELEX: 66141 PENNIE  
 ; INFORMATION FOR SEQ ID NO: 5:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3321 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: DNA (genomic)  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 156..1782  
 ; US-08-484-438-5

Query Match 2.7%; Score 34.6; DB 1; Length 3321;  
 Best Local Similarity 51.6%; Pred. No. 2.2;  
 Matches 79; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 128 GCAGAAATGAGAGGAGGAGCTGTGGACTGGTTACTTTTATATACTTACCTTAAAGAC 187

Db 1378 GAAAAAATGGAGAGCCCTTCAAGCATTTGGATAATCCCAATATACAAATGCATCCAATGGTC 1437  
 QY 188 AAAACAAGAAAGTGGAGAGACTGGTTAGAGTACCTGTACTAGACTCTTACAACACTAGAA 247  
 Db 1438 CACCCAGGCCGAGGATGAGTATGTGAATGAGCCACTGTACTCTCAACACCTTTGGCAACA 1497  
 QY 248 GCTGGAGGAAGAGTGAGCAACTAATGAATGACA 280  
 Db 1498 CCTTGGGAAAAAGCTGAGTACCTGAAGAAACA 1530

Search completed: October 14, 2004, 05:27:54  
 Job time : 112.854 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2004, 00:31:21 : Search time 67.2152 Seconds  
(without alignments)  
1694.809 Million cell updates/sec

Title: US-10-790-589-2

Perfect score: 1927

Sequence: 1 MTAKPLRTVLSLLFFALSGV.....KNHIYQAFHKLRYGCK 354

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Lasting first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1927	100.0	354	9	US-09-949-434-2
2	1927	100.0	354	12	US-10-240-709-2
3	1927	100.0	354	16	US-10-790-589-2
4	1927	100.0	354	16	US-10-670-863-1
5	1293.5	67.1	357	9	US-09-949-434-2
6	1293.5	67.1	357	12	US-10-240-709-2
7	1293.5	67.1	357	16	US-10-790-589-2
8	1286.5	66.8	361	16	US-10-670-863-3
9	596.5	31.0	360	15	US-10-408-167A-2
10	557.5	28.9	348	16	US-10-408-765A-1164
11	111.5	5.8	818	16	US-10-437-963-145908
12	111.5	5.8	832	16	US-10-437-963-145910
13	101.5	5.3	475	9	US-09-729-674-142
14	101.5	5.3	475	9	US-09-817-774-26
15	101.5	5.3	475	16	US-10-429-160-32

16	97.5	5.1	1886	12	US-10-147-299A-6	Sequence 6, Appli
17	97.5	5.1	2499	16	US-10-706-424-4	Sequence 4, Appli
18	94	4.9	1385	9	US-09-738-363-2	Sequence 2, Appli
19	94	4.9	1385	15	US-10-633-023-2	Sequence 2, Appli
20	94	4.9	2165	10	US-09-923-070A-29	Sequence 29, Appli
21	94	4.9	2165	10	US-09-923-070A-30	Sequence 30, Appli
22	94	4.9	2165	10	US-09-161-122-29	Sequence 29, Appli
23	94	4.9	2165	10	US-09-161-122-30	Sequence 30, Appli
24	94	4.9	2165	16	US-10-628	Sequence 393, App
25	93.5	4.9	361	9	US-09-925-300-1693	Sequence 1693, Ap
26	89	4.6	267	13	US-10-000-954-3	Sequence 3, Appli
27	88.5	4.6	2771	9	US-09-808-602-82	Sequence 82, Appli
28	88.5	4.6	2771	10	US-09-800-198-70	Sequence 70, Appli
29	87.5	4.5	780	12	US-10-282-122A-45975	Sequence 45975, A
30	87.5	4.5	3868	16	US-10-461-194-103	Sequence 103, App
31	87	4.5	607	12	US-10-425-114-67461	Sequence 67461, A
32	87	4.5	778	16	US-10-437-963-163162	Sequence 163162,
33	87	4.5	900	12	US-10-282-122A-63553	Sequence 63553, A
34	87	4.5	964	12	US-10-042-865-55	Sequence 55, Appli
35	87	4.5	1045	12	US-10-042-865-54	Sequence 54, Appli
36	87	4.5	1045	12	US-10-029-020-54	Sequence 54, Appli
37	87	4.5	2613	16	US-10-038-854-42	Sequence 42, Appli
38	87	4.5	2628	16	US-10-038-854-40	Sequence 40, Appli
39	87	4.5	2721	16	US-10-038-854-38	Sequence 38, Appli
40	87	4.5	2725	16	US-10-038-854-36	Sequence 36, Appli
41	86	4.5	424	12	US-10-282-122A-55910	Sequence 55910, A
42	86	4.5	604	16	US-10-648-152-2	Sequence 2, Appli
43	86	4.5	618	16	US-10-648-152-37	Sequence 37, Appli
44	86	4.5	624	16	US-10-648-152-29	Sequence 29, Appli
45	86	4.5	1005	12	US-10-425-114-63864	Sequence 63864, A

## ALIGNMENTS

### RESULT 1

US-09-949-434-2  
; Sequence 2, Application US/09949434  
; Patent No. US20020028495A1  
; GENERAL INFORMATION:  
; APPLICANT: Eastman, Alan R.  
; APPLICANT: Krieser, Ronald J.  
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND cDNAs  
; FILE REFERENCE: DC-0137  
; CURRENT APPLICATION NUMBER: US/09/949,434  
; CURRENT FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: 09/574,942  
; PRIOR FILING DATE: 2000-05-19  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; TYPE: PRT  
; ORGANISM: Mus sp.  
US-09-949-434-2

Query Match	100.0%	Score 1927;	DB 9;	Length 354;
Best Local Similarity	100.0%	Pred. NO. 7.6e-189;		
Matches 354;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MTAKPLRTVLSLLFFALSGVLTGTPISCRNEGEADVDFYKLPKRTGKASEAGLQYL	60	
Db	1	MTAKPLRTVLSLLFFALSGVLTGTPISCRNEGEADVDFYKLPKRTGKASEAGLQYL	60	
QY	61	YLDSTRQTNKSLYLINSTRSALGRTLQHLPTHSTNDTAYLIYNDGVPGSVNYRQYK	120	
Db	61	YLDSTRQTNKSLYLINSTRSALGRTLQHLPTHSTNDTAYLIYNDGVPGSVNYRQYK	120	
QY	121	HAKGLLWNRTOGFWLIHSVPKPPVHGVEYTSRRYQGTGTCITFGYSQFEEDFOLL	180	
Db	121	HAKGLLWNRTOGFWLIHSVPKPPVHGVEYTSRRYQGTGTCITFGYSQFEEDFOLL	180	
QY	181	VLQPNYSFCFISTFWKLIYMPRCANSSSLKIPVRYLAELHLSAQGLNFVHFPSFYT	240	

```

181 VLOPNISCTPSTFHHKLIYMPMCANSSSLKIPVRYLAELHSAQGLNFVHFAKSSFT 240
241 DDIFTGIAOKLTHLLAQTWKKQELPNSCLPYHVYNIKSIQVTSKYSFSSRDHSHK 300
241 DDIFTGIAOKLTHLLAQTWKKQELPNSCLPYHVYNIKSIQVTSKYSFSSRDHSHK 300
301 WCVSIKGSANRWTCIGDLNLSHQAALRGGGFICTKNHYIYQAFHKLRYGFCK 354
301 WCVSIKGSANRWTCIGDLNLSHQAALRGGGFICTKNHYIYQAFHKLRYGFCK 354

```

RESULT 2

```

US-10-790-709-2
; Sequence 2, Application US/10240709
; Publication No. US20030212023A1
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Krieser, Ronald J.
; APPLICANT: Trustees of Dartmouth College
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCE: DC-0154
; CURRENT APPLICATION NUMBER: US/10/240,709
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-790-709-2

```

```

Query Match 100.0%; Score 1927; DB 12; Length 354;
Best Local Similarity 100.0%; Pred. No. 7.6e-189;
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTAKPLRTVLSLFFALSGLVGTPEISCRNEYGEAVDWFIFYKLPKRTSKASEAGLOYL 60
Db 1 MTAKPLRTVLSLFFALSGLVGTPEISCRNEYGEAVDWFIFYKLPKRTSKASEAGLOYL 60

Qy 61 YLDSTRTQWTKSLYLINSTSALGRTLOHLYDTHNSTNDTAYLYNDGVPGSVNSRYG 120
Db 61 YLDSTRTQWTKSLYLINSTSALGRTLOHLYDTHNSTNDTAYLYNDGVPGSVNSRYG 120

Qy 121 HAKGLLVNRTQGFLLIHSVPKPPVHGYEYPTSGRRYGTGICITFGYSQFEIDFQLL 180
Db 121 HAKGLLVNRTQGFLLIHSVPKPPVHGYEYPTSGRRYGTGICITFGYSQFEIDFQLL 180

Qy 181 VLOPNISCTPSTFHHKLIYMPMCANSSSLKIPVRYLAELHSAQGLNFVHFAKSSFT 240
Db 181 VLOPNISCTPSTFHHKLIYMPMCANSSSLKIPVRYLAELHSAQGLNFVHFAKSSFT 240

Qy 241 DDIFTGIAOKLTHLLAQTWKKQELPNSCLPYHVYNIKSIQVTSKYSFSSRDHSHK 300
Db 241 DDIFTGIAOKLTHLLAQTWKKQELPNSCLPYHVYNIKSIQVTSKYSFSSRDHSHK 300

Qy 301 WCVSIKGSANRWTCIGDLNLSHQAALRGGGFICTKNHYIYQAFHKLRYGFCK 354
Db 301 WCVSIKGSANRWTCIGDLNLSHQAALRGGGFICTKNHYIYQAFHKLRYGFCK 354

```

RESULT 3

```

US-10-790-589-2
; Sequence 2, Application US/10790589
; Publication No. US20040142376A1
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Krieser, Ronald J.
; APPLICANT: Trustees of Dartmouth College

```

```

; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCE: DC-0154
; CURRENT APPLICATION NUMBER: US/10/790,589
; CURRENT FILING DATE: 2004-03-01
; PRIOR APPLICATION NUMBER: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-790-589-2

```

```

Query Match 100.0%; Score 1927; DB 16; Length 354;
Best Local Similarity 100.0%; Pred. No. 7.6e-189;
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTAKPLRTVLSLFFALSGLVGTPEISCRNEYGEAVDWFIFYKLPKRTSKASEAGLOYL 60
Db 1 MTAKPLRTVLSLFFALSGLVGTPEISCRNEYGEAVDWFIFYKLPKRTSKASEAGLOYL 60

Qy 61 YLDSTRTQWTKSLYLINSTSALGRTLOHLYDTHNSTNDTAYLYNDGVPGSVNSRYG 120
Db 61 YLDSTRTQWTKSLYLINSTSALGRTLOHLYDTHNSTNDTAYLYNDGVPGSVNSRYG 120

Qy 121 HAKGLLVNRTQGFLLIHSVPKPPVHGYEYPTSGRRYGTGICITFGYSQFEIDFQLL 180
Db 121 HAKGLLVNRTQGFLLIHSVPKPPVHGYEYPTSGRRYGTGICITFGYSQFEIDFQLL 180

Qy 181 VLOPNISCTPSTFHHKLIYMPMCANSSSLKIPVRYLAELHSAQGLNFVHFAKSSFT 240
Db 181 VLOPNISCTPSTFHHKLIYMPMCANSSSLKIPVRYLAELHSAQGLNFVHFAKSSFT 240

Qy 241 DDIFTGIAOKLTHLLAQTWKKQELPNSCLPYHVYNIKSIQVTSKYSFSSRDHSHK 300
Db 241 DDIFTGIAOKLTHLLAQTWKKQELPNSCLPYHVYNIKSIQVTSKYSFSSRDHSHK 300

Qy 301 WCVSIKGSANRWTCIGDLNLSHQAALRGGGFICTKNHYIYQAFHKLRYGFCK 354
Db 301 WCVSIKGSANRWTCIGDLNLSHQAALRGGGFICTKNHYIYQAFHKLRYGFCK 354

```

RESULT 4

```

US-10-670-863-1
; Sequence 1, Application US/10670863
; Publication No. US20040157239A1
; GENERAL INFORMATION:
; APPLICANT: Tanuma, Sei-ichi
; APPLICANT: Shiokawa, Daisuke
; TITLE OF INVENTION: Novel Deoxyribonuclease, Gene Encoding Thereof and Use Thereof
; FILE REFERENCE: 224589
; CURRENT APPLICATION NUMBER: US/10/670,863
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: JP 11-230870
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: US 09/807,784
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 15
; SEQ ID NO 1
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-670-863-1

```

```

Query Match 100.0%; Score 1927; DB 16; Length 354;
Best Local Similarity 100.0%; Pred. No. 7.6e-189;
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTAKPLRTVLSLFFALSGLVGTPEISCRNEYGEAVDWFIFYKLPKRTSKASEAGLOYL 60

```

Db 1 MTAKPLRTVLSLFFALSGVLGTPETSCRENEYGEAVDWFIFYKLPKRTSKASEEAGLOYL 60  
 QY 61 YLDSTQTNKSLYLINSTRSALGRTLOHLYDTHNSTNDTAYLIYNDGVPGSVNYSRQY 120  
 Db 61 YLDSTQTNKSLYLINSTRSALGRTLOHLYDTHNSTNDTAYLIYNDGVPGSVNYSRQY 120  
 QY 121 HAKGLLVNRTQGFMLIHSPKPPVHGYEPTSGRRYGTGTCITFGYSQSEEDIFQLL 180  
 Db 121 HAKGLLVNRTQGFMLIHSPKPPVHGYEPTSGRRYGTGTCITFGYSQSEEDIFQLL 180  
 QY 181 VLQPNYSCFIPSTFHWKLIYMPRMCANSSSLKIPVRYLAELHSAQGLNPFVHFAKSSFTY 240  
 Db 181 VLQPNYSCFIPSTFHWKLIYMPRMCANSSSLKIPVRYLAELHSAQGLNPFVHFAKSSFTY 240  
 QY 241 DDIPTGWIQAUKLTHLLAOTWOKKQELPNSCLPYPVYVNIKSIIGVTSKYSFSSRODHSK 300  
 Db 241 DDIPTGWIQAUKLTHLLAOTWOKKQELPNSCLPYPVYVNIKSIIGVTSKYSFSSRODHSK 300  
 QY 301 WCVSINGSANRWTICIGDLNRSLHQALRGCGGFTCKNHYIYQAFHKLYLYRGFCK 354  
 Db 301 WCVSINGSANRWTICIGDLNRSLHQALRGCGGFTCKNHYIYQAFHKLYLYRGFCK 354

## RESULT 5

US-09-949-434-4  
 ; Sequence 4, Application US/09949434  
 ; Patent No. US20020028495A1

GENERAL INFORMATION:  
 ; APPLICANT: Eastman, Alan R.

; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS

; FILE REFERENCE: DC-0137

; CURRENT APPLICATION NUMBER: US/09/949,434

; PRIOR FILING DATE: 2001-09-07

; PRIOR FILING DATE: 2000-05-19

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 357

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-949-434-4

Query Match 67.1%; Score 1293.5; DB 9; Length 357;  
 Best Local Similarity 66.4%; Pred. No. 8.4e-124;  
 Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

QY 1 MTAKPLRTVLSLFFALSGVLGTPETSCRENEYGEAVDWFIFYKLPKRTSKASEEAGLOYL 60  
 Db 1 MWARLRTSFALLFLGLFGLGAATISCRNEGKAVDWFIFYKLPKRTSKASEEAGLOYL 60  
 QY 61 YLDSTQTNKSLYLINSTRSALGRTLOHLYDTHNSTNDTAYLIYNDGVPGSVNYSRQY 119  
 Db 61 YLDSTTRSWRKSEQLMNDTKSVLGRTLQOLYEAYASKSNNTAYLIYNDGVPKPVNYSRKY 120  
 QY 120 GHAKGLLVNRTQGFMLIHSPKPPV--HGVEYPTSGRRYGTGTCITFGYSQSEEDIF 177  
 Db 121 GHTKGLLLNVRVQGFMLIHSPFPPEEGYDPTGRRNGSGICITFKYNQYEADS 180  
 QY 178 QLLVLPQNIYSCFIPSTFHWKLIYMPRMCANSSSLKIPVRYLAELHSAQGLNPFVHFAKSS 237  
 Db 181 QLLVCNPNVYSCSIPATFHQELIHPQLCTRASSSEIPGELLTTTQSAGQKFLHFAKSD 240  
 QY 238 FTYDDIFGWIQAUKLTHLLAOTWOKKQELPNSCLPYPVYVNIKSIIGVTSKYSFSSROD 297  
 Db 241 SFLLDDIFAANWAQRLKTHLLTETWQRKQELPNSCLPYPVYVNIKAIKLSRHSYFSSYQD 300  
 QY 298 HSKWCVSINGSANRWTICIGDLNRSLHQALRGCGGFTCKNHYIYQAFHKLYLYRGFCK 354  
 Db 301 HAKWCISQKGTNRWTICIGDLNRSHPQAFRSGGFTCTQNWQIYQAFQGLVLYESCK 357

## RESULT 6

US-10-240-709-4

; Sequence 4, Application US/10240709  
 ; Publication No. US20030212023A1

GENERAL INFORMATION:  
 ; APPLICANT: Eastman, Alan R.

; APPLICANT: Krieser, Ronald J.

; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS

; FILE REFERENCE: DC-0154

; CURRENT APPLICATION NUMBER: US/10/240,709

; CURRENT FILING DATE: 2002-10-01

; PRIOR APPLICATION NUMBER: 09/541,840

; PRIOR FILING DATE: 2000-04-03

; PRIOR APPLICATION NUMBER: 09/574,942

; PRIOR FILING DATE: 2000-05-19

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 357

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-240-709-4

Query Match 67.1%; Score 1293.5; DB 12; Length 357;  
 Best Local Similarity 66.4%; Pred. No. 8.4e-124;  
 Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

QY 1 MTAKPLRTVLSLFFALSGVLGTPETSCRENEYGEAVDWFIFYKLPKRTSKASEEAGLOYL 60  
 Db 1 MWARLRTSFALLFLGLFGLGAATISCRNEGKAVDWFIFYKLPKRTSKASEEAGLOYL 60  
 QY 61 YLDSTQTNKSLYLINSTRSALGRTLOHLYDTHNSTNDTAYLIYNDGVPGSVNYSRQY 119  
 Db 61 YLDSTTRSWRKSEQLMNDTKSVLGRTLQOLYEAYASKSNNTAYLIYNDGVPKPVNYSRKY 120  
 QY 120 GHAKGLLVNRTQGFMLIHSPKPPV--HGVEYPTSGRRYGTGTCITFGYSQSEEDIF 177  
 Db 121 GHTKGLLLNVRVQGFMLIHSPFPPEEGYDPTGRRNGSGICITFKYNQYEADS 180  
 QY 178 QLLVLPQNIYSCFIPSTFHWKLIYMPRMCANSSSLKIPVRYLAELHSAQGLNPFVHFAKSS 237  
 Db 181 QLLVCNPNVYSCSIPATFHQELIHPQLCTRASSSEIPGELLTTTQSAGQKFLHFAKSD 240  
 QY 238 FTYDDIFGWIQAUKLTHLLAOTWOKKQELPNSCLPYPVYVNIKSIIGVTSKYSFSSROD 297  
 Db 241 SFLLDDIFAANWAQRLKTHLLTETWQRKQELPNSCLPYPVYVNIKAIKLSRHSYFSSYQD 300  
 QY 298 HSKWCVSINGSANRWTICIGDLNRSLHQALRGCGGFTCKNHYIYQAFHKLYLYRGFCK 354  
 Db 301 HAKWCISQKGTNRWTICIGDLNRSHPQAFRSGGFTCTQNWQIYQAFQGLVLYESCK 357

## RESULT 7

US-10-790-589-4

; Sequence 4, Application US/10790589  
 ; Publication No. US20040142376A1

GENERAL INFORMATION:  
 ; APPLICANT: Eastman, Alan R.

; APPLICANT: Krieser, Ronald J.

; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS

; FILE REFERENCE: DC-0154

; CURRENT APPLICATION NUMBER: US/10/790,589

; CURRENT FILING DATE: 2004-03-01

; PRIOR APPLICATION NUMBER: 09/541,840

; PRIOR FILING DATE: 2000-04-03

; PRIOR APPLICATION NUMBER: 09/574,942

; PRIOR FILING DATE: 2000-05-19

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 357

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-790-589-4

Query Match      67.1%; Score 1293.5; DB 16; Length 357;
Best Local Similarity 66.4%; Pred. No. 8.4e-124;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

Qy 1 MTAKPLTVLSLFFALSGVLGTPETISCRNEYGAVDWFIYKLPKRTSKASEAGLQYL 60
Db 1 MMARLLRTSFAALLFLGLFGLVGAATISCRNEBKAVDWFTFYKLPKQNKESGETGLEYL 60

Qy 61 YLDSTRCTWNSLYLINSTSALGRTOHLIYDTHNS--TNDTAYLIYNDGVPGSVNYSGROY 119
Db 61 YLDSTRCTWNSLYLINSTSALGRTOHLIYDTHNS--TNDTAYLIYNDGVPGSVNYSGROY 119

Qy 120 GHAKGLIYVNRTOGFWLIHSPVKPPV--HGVEYPTSGRRYQGTGICITFGYSQFEEDF 177
Db 120 GHAKGLIYVNRTOGFWLIHSPVKPPV--HGVEYPTSGRRYQGTGICITFGYSQFEEDF 177

Qy 178 QLVLPQNIYSCPTSTFHWKLIYMPRMCAVNSSLKIPVRYLAELHSAQGLNFVHFAKSS 237
Db 178 QLVLPQNIYSCPTSTFHWKLIYMPRMCAVNSSLKIPVRYLAELHSAQGLNFVHFAKSS 237

Qy 181 QLLVCNPNVYSCIPATFHQELIHPOLCTRASSSEIPGRLLTTLQSAQOQKFLHFAKSD 240
Db 181 QLLVCNPNVYSCIPATFHQELIHPOLCTRASSSEIPGRLLTTLQSAQOQKFLHFAKSD 240

Qy 238 FYTDDIFTGIAOKLTHLLAQTWOKKQELPNSCLPYHYVNIKSTIGVTSKSYFSSROD 297
Db 238 FYTDDIFTGIAOKLTHLLAQTWOKKQELPNSCLPYHYVNIKSTIGVTSKSYFSSROD 297

Qy 241 SFLLDDIPAAWMAQKLTHTLTTETWOKRQELPNSCLPYHYVNIKSTIGVTSKSYFSSROD 300
Db 241 SFLLDDIPAAWMAQKLTHTLTTETWOKRQELPNSCLPYHYVNIKSTIGVTSKSYFSSROD 300

Qy 298 HSKWCYSIKGSANRWTCIGDLNLSLHQAALRGGGFICTKKNHYIYQAFHKLRYLRYGFK 354
Db 298 HSKWCYSIKGSANRWTCIGDLNLSLHQAALRGGGFICTKKNHYIYQAFHKLRYLRYGFK 354

Qy 301 HAKWCISQKTKNRTKTCIGDLNRSHPHQAQFSGGFICTQNWQIYQAFQGLVLYESCK 357
Db 301 HAKWCISQKTKNRTKTCIGDLNRSHPHQAQFSGGFICTQNWQIYQAFQGLVLYESCK 357

RESULT 8
US-10-670-863-3
; Sequence 3, Application US/10670863
; Publication No. US20040157239A1
; GENERAL INFORMATION:
; APPLICANT: Tanuma, Sei-ichi
; APPLICANT: Shiohara, Daisuke
; TITLE OF INVENTION: Novel Deoxyribonuclease, Gene Encoding Thereof and Use Thereof
; FILE REFERENCE: 224589
; CURRENT APPLICATION NUMBER: US/10/670,863
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: JP 11-230870
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: US 09/807,784
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 15
; SEQ ID NO 3
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-670-863-3

Query Match      66.8%; Score 1286.5; DB 16; Length 361;
Best Local Similarity 66.1%; Pred. No. 4.4e-123;
Matches 236; Conservative 45; Mismatches 73; Indels 3; Gaps 2;

Qy 1 MTAKPLTVLSLFFALSGVLGTPETISCRNEYGAVDWFIYKLPKRTSKASEAGLQYL 60
Db 5 MMARLLRTSFAALLFLGLFGLVGAATISCRNEBKAVDWFTFYKLPKQNKESGETGLEYL 64

Qy 61 YLDSTRCTWNSLYLINSTSALGRTOHLIYDTHNS--TNDTAYLIYNDGVPGSVNYSGROY 119
Db 65 YLDSTRCTWNSLYLINSTSALGRTOHLIYDTHNS--TNDTAYLIYNDGVPGSVNYSGROY 124

Qy 120 GHAKGLIYVNRTOGFWLIHSPVKPPV--HGVEYPTSGRRYQGTGICITFGYSQFEEDF 177
Db 125 GHAKGLIYVNRTOGFWLIHSPVKPPV--HGVEYPTSGRRYQGTGICITFGYSQFEEDF 177

Qy 178 QLVLPQNIYSCPTSTFHWKLIYMPRMCAVNSSLKIPVRYLAELHSAQGLNFVHFAKSS 237
Db 178 QLVLPQNIYSCPTSTFHWKLIYMPRMCAVNSSLKIPVRYLAELHSAQGLNFVHFAKSS 237

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-790-589-4

Query Match      67.1%; Score 1293.5; DB 16; Length 357;
Best Local Similarity 66.4%; Pred. No. 8.4e-124;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

Qy 1 MTAKPLTVLSLFFALSGVLGTPETISCRNEYGAVDWFIYKLPKRTSKASEAGLQYL 60
Db 1 MMARLLRTSFAALLFLGLFGLVGAATISCRNEBKAVDWFTFYKLPKQNKESGETGLEYL 60

Qy 61 YLDSTRCTWNSLYLINSTSALGRTOHLIYDTHNS--TNDTAYLIYNDGVPGSVNYSGROY 119
Db 61 YLDSTRCTWNSLYLINSTSALGRTOHLIYDTHNS--TNDTAYLIYNDGVPGSVNYSGROY 119

Qy 120 GHAKGLIYVNRTOGFWLIHSPVKPPV--HGVEYPTSGRRYQGTGICITFGYSQFEEDF 177
Db 120 GHAKGLIYVNRTOGFWLIHSPVKPPV--HGVEYPTSGRRYQGTGICITFGYSQFEEDF 177

Qy 178 QLVLPQNIYSCPTSTFHWKLIYMPRMCAVNSSLKIPVRYLAELHSAQGLNFVHFAKSS 237
Db 178 QLVLPQNIYSCPTSTFHWKLIYMPRMCAVNSSLKIPVRYLAELHSAQGLNFVHFAKSS 237

Qy 181 QLLVCNPNVYSCIPATFHQELIHPOLCTRASSSEIPGRLLTTLQSAQOQKFLHFAKSD 240
Db 181 QLLVCNPNVYSCIPATFHQELIHPOLCTRASSSEIPGRLLTTLQSAQOQKFLHFAKSD 240

Qy 238 FYTDDIFTGIAOKLTHLLAQTWOKKQELPNSCLPYHYVNIKSTIGVTSKSYFSSROD 297
Db 238 FYTDDIFTGIAOKLTHLLAQTWOKKQELPNSCLPYHYVNIKSTIGVTSKSYFSSROD 297

Qy 241 SFLLDDIPAAWMAQKLTHTLTTETWOKRQELPNSCLPYHYVNIKSTIGVTSKSYFSSROD 300
Db 241 SFLLDDIPAAWMAQKLTHTLTTETWOKRQELPNSCLPYHYVNIKSTIGVTSKSYFSSROD 300

Qy 298 HSKWCYSIKGSANRWTCIGDLNLSLHQAALRGGGFICTKKNHYIYQAFHKLRYLRYGFK 354
Db 298 HSKWCYSIKGSANRWTCIGDLNLSLHQAALRGGGFICTKKNHYIYQAFHKLRYLRYGFK 354

Qy 301 HAKWCISQKTKNRTKTCIGDLNRSHPHQAQFSGGFICTQNWQIYQAFQGLVLYESCK 357
Db 301 HAKWCISQKTKNRTKTCIGDLNRSHPHQAQFSGGFICTQNWQIYQAFQGLVLYESCK 357

RESULT 9
US-10-408-167A-2
; Sequence 2, Application US/10408167A
; Publication No. US20030219428A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Baron, Will F.
; TITLE OF INVENTION: Human DNase II
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/408,167A
; FILING DATE: 04-Apr-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/861034
; FILING DATE: 18-May-2001
; APPLICATION NUMBER: 08/639294
; FILING DATE: 25-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Evans, David W.
; REGISTRATION NUMBER: NONE
; REFERENCE/DOCKET NUMBER: P1024D1C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1739
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-408-167A-2

Query Match      31.0%; Score 596.5; DB 15; Length 360;
Best Local Similarity 38.0%; Pred. No. 3e-52;
Matches 132; Conservative 56; Mismatches 144; Indels 15; Gaps 7;

Qy 17 LSGVLTGP--EISCRNEYGAVDWFIYKLPK-RTSKASEAGLQYLYLDSTRQTNKSL 73
Db 6 LAALLCVFAGALTCTGDSGQVDFVYVYKLPALRGSGEAAQGRQYKLYLDESSGGRDGR 65

Qy 74 YLINSTSALGRTOHLIYDTHNS--TNDTAYLIYND--GVPGSVNYSGROYGHAKGLIYVNR 131
Db 66 ALINSPGAVGRSLQPLY--RSNTSQLAFLLYNDQPPQPSKAQDSSWRGHTKGVLLIHD 123

Qy 132 QGFVLIHSPVKPP--VHGVEYPTSGRRYQGTGICITFGYSQFEEDFOLLVLPQNIYS 188
Db 124 GGFVLIHSPVNFPPASSAAVSWPHSACTYQITLLCVSFPPAQFSKMGKQLTYTPWYN 183

Qy 189 CFIPSTPHWKLIYMPRMCAVNSSLKIPVRYLAELHSAQGLNFVHFAKSSFTDIDFTGWI 248

```



Db 184 YQLEGFAQEFPPDENLVKHHVYQEBWSSITLTSQAGAVFQSFQKFKFGDDLYSGWL 243  
QY 249 AQKLTHTLLAQTWQKKQELPNSCSLPYHYNTKISGTVTSKS--YFSSRODHKWCVCVK 306  
Db 244 AAALGNLQVQFVHKTVGILLPNSCSDIWOVLNVQIAFPAGPAGESFNSTEDHKKWCVCVK 303  
QY 307 GSNARWTCIGDLNRSLHQALRGGGFICTKNHNTYQAFHKLILRYGFC 353  
Db 304 GP---WTCVGMNRNQEGEGGTLCAQLPALWKAQFPLVKNYQPC 347

## RESULT 10

US-10-408-765A-1164  
; Sequence 1164, Application US/10408765A  
; Publication No. US20040101874A1  
; GENERAL INFORMATION:  
; APPLICANT: Ghosh, Soumitra S.  
; APPLICANT: Fahy, Eoin D.  
; APPLICANT: Zhang, Bing  
; APPLICANT: Gibson, Bradford W.  
; APPLICANT: Taylor, Steven W.  
; APPLICANT: Glenn, Gary M.  
; APPLICANT: Warnock, Dale E.  
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
; FILE REFERENCE: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME  
; CURRENT APPLICATION NUMBER: US/10/408,765A  
; CURRENT FILING DATE: 2003-04-04  
; NUMBER OF SEQ ID NOS: 3077  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1164  
; LENGTH: 348  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-408-765A-1164

Query Match 28.9%; Score 557.5; DB 16; Length 348;  
Best Local Similarity 37.0%; Pred. No. 2.9e-48;  
Matches 129; Conservative 52; Mismatches 137; Indels 31; Gaps 9;

QY 17 LSGVLTGP--EISCRNEYGEAVDWFIFKLPK-RTSKASEAGLQYLVDSTRQTNWKS 73  
Db 6 LAALLCVPAALTCYGDSDQPDVFWYKLPALRGSGEAAQGLQYKYLDESSGGWRDGR 65  
QY 74 YLINSRTRALGTLQHLTDHNSNDTAYLIYND--GVPGSVNYSGRYQGHAKGLLVNRT 131  
Db 66 ALINSEGAAGVRSLOPLY--RSNTSQLAEFLYNDQPPQSKAQDSMRGHTKGVLILDHD 123  
QY 132 QGFWLHISVPKPPP--VHGVEYPTSGRRYGOTGIC--ITFGYSQFEEDFOLLVLQPMI 186  
Db 124 GGFVLVHVSFVNFPPPPASSAAVSNPHSACTYGTLLCKQLTYTY-----FWV 169  
QY 187 YSCFTPTSTHFWKLIYNRMCAANSLSKIPVRYLAELHSAQGLNFVHFPAKSSFYTDIFTG 246  
Db 170 YNYQLEGIFAQEFPPDENLVKHHVYQEBWSSITLTSQAGAVFQSFQKFKFGDDLYSG 229  
QY 247 WIAQKLTHTLLAQTWQKKQELPNSCSLPYHYNTKISGTVTSKS--YFSSRODHKWCVCVK 304  
Db 230 WIAAALGNLQVQFVHKTVGILLPNSCSDIWOVLNVQIAFPAGPAGESFNSTEDHKKWCVCVK 289  
QY 305 IYGSANRWTCIGDLNRSLHQALRGGGFICTKNHNTYQAFHKLILRYGFC 353  
Db 290 PKGP---WTCVGMNRNQEGEGGTLCAQLPALWKAQFPLVKNYQPC 335

## RESULT 11

US-10-437-963-145908  
; Sequence 145908, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: Plants and Uses Thereof for Plant Improvement  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 145908  
; LENGTH: 818  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_46584C.1.1.pap  
US-10-437-963-145908

Query Match 5.8%; Score 111.5; DB 16; Length 818;  
Best Local Similarity 20.8%; Pred. No. 0.06;  
Matches 59; Conservative 33; Mismatches 111; Indels 81; Gaps 12;  
QY 23 TPEISCRNEYGEAVDWFIFKLPKRTSKASEAGLQYLVDSTRQTNWKSILYNSTRGA 82  
Db 250 TGFSCSGDPSSDLQVFWH-----GTRPYRSIVLDS---VW-----VSGKA 289  
QY 83 LGRTLQHLTDHNSNDTAYLIY--NDGVPGSVNYSR---QYGHAKGLLVNRTQGFWLI 137  
Db 290 YGSSYSPMYQTVNTQDEFYIYVTSQSP-----YMRIMLDYTGFRLLSWNVNSSWAI 345  
QY 138 HSNVPPPPVHGVEYPTSGRRYGOTGICITFGYSQFEEDFOLLVLQPMIYSCFIPSTFEW 197  
Db 346 YS-----QRDAAGICDPYGSQGPFGYCDFTSV-----IPRCQCPDGF- 384  
QY 198 KLIYMRMCANSSSLKIPVRYLAELHSAQGLNF-----VHFPAKSSFYTDIFGTIAQKL 252  
Db 385 -----PNSNSSSGCRRKQQLRCGEGNHFTWPMGKLPDKPFYVD----- 425  
QY 253 KTHLLAQTWQKKQELPNSCSLPYHYVY-NIKSIGVTSKSYFSSR 295  
Db 426 -----RSFECAACSRNCSTAYTNLTITGSPGTTASQSR 463

## RESULT 12

US-10-437-963-145910  
; Sequence 145910, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: Plants and Uses Thereof for Plant Improvement  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 145910  
; LENGTH: 832  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_46586C.1.1.pap  
US-10-437-963-145910

Query Match

5.8%; Score 111.5; DB 16; Length 832;

Best Local Similarity 20.8%; Pred. No. 0.061;  
Matches 59; Conservative 33; Mismatches 111; Indels 81; Gaps 12;  
QY 23 TPBISCRNEYGEAVDMFIFYLKPKRTSKASEEAGLOLYLDSTROTWNKSLYLINSTRSA 82  
Db 264 TGFSCSGDSSDLQVFWH-----GTPKPYRSIVLDS--VM-----VSGKA 303  
QY 83 LGTLQHLNTHSTNDTAYLIY--NDGVPGSVNSYR--QYGHAKGLLVNWKRTQGFWLI 137  
Db 304 YGSSTSFMYQTYNTQDEFVIYITSDGSP-----YMRIMLDYTGTFRLLSWNNSSSWAI 359  
QY 138 HSPVKPPVHGVEYPTSGRRYGTGICITFGYSOFEEIDFQLLVLPQNIYSCFIPSTFWH 197  
Db 360 YS-----QRPAAIGCDPVGSGPGYCDFTSV-----IPRCQCPDGE- 398  
QY 198 KLIYMPRCANSSSLKIPVRYLAELHSAQGLNF-----VHFPAKSFYTDITFGWIAQKL 252  
Db 399 -----PENGSSSGCRRKQQLRCGEGHFMFMKLPDKFFVQD----- 439  
QY 253 KTHLAQTWQKKQELPSNCSLPYHYV-NIKSIGVTSKSYESSR 295  
Db 440 -----RSFECAACSRNCSTAYATNLTITGSPGTTASQSR 477

RESULT 13  
US-09-729-674-142  
; Sequence 142, Application US/09729674  
; Patent No. US20010039335A1  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Kenneth  
; APPLICANT: McCoy, John M.  
; APPLICANT: LaValle, Edward R.  
; APPLICANT: Collins-Racie, Lisa A.  
; APPLICANT: Evans, Cheryl  
; APPLICANT: Merberg, David  
; APPLICANT: Treacy, Maurice  
; APPLICANT: Agostino, Michael J.  
; APPLICANT: Steining, Robert J.  
; APPLICANT: Spaulding, Vikki  
; APPLICANT: Wong, Gordon G.  
; APPLICANT: Clark, Hilary  
; APPLICANT: Fechtel, Kim  
; APPLICANT: Genetics Institute, Inc.  
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM  
; FILE REFERENCE: 6055-64X  
; CURRENT APPLICATION NUMBER: US/09/729,674  
; PRIOR FILING DATE: 2000-12-04  
; PRIOR APPLICATION NUMBER: 09/539,330  
; PRIOR FILING DATE: 2000-03-30  
; NUMBER OF SEQ ID NOS: 283  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 142  
; LENGTH: 475  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-729-674-142

Query Match 5.3%; Score 101.5; DB 9; Length 475;  
Best Local Similarity 21.4%; Pred. No. 0.29;  
Matches 100; Conservative 48; Mismatches 147; Indels 173; Gaps 23;  
QY 6 LRTVLSLLFFA----LSGVLTGPISCRNEYGEAVDMF-----IPYKLPKRTSKASE 53  
Db 40 LASVIFLLFPFIVYFIMACDQVSCALT-GPVVDIVTGHARLSDIWAKTPPTRKAAQ 98  
QY 54 BAGL-----QYLYLDSRTQWNKSL-----YLNSTRSALGRTHQ 89  
Db 99 LYTLLVTFQVLLYTSLPDFCHKFLPGYVGGIOGAVTPAGVNVKQYINGLQAWLLTHLLW 158  
QY 90 LYDTHNSTNDTAYLIYNDGVPGSVNSYRQYGHAKGLLVNWKRTQGFWLIHSPKPPVHG 149  
Db 159 FANAHLWSFSTIIFDWNIP-----LLWCAN-----ILGAVSTFANVKGY 200

QY 150 EYPTSGRRYGTGICITFGYSOFEEIDFQLLVLPQNIYSCF-----IPSTFWHKLIIY 201  
Db 201 FFPTSARDCKFTG---NFFYNMVGIEF-----NPRIGKWFDFKLFNGREGIVAWTLIN 252  
QY 202 MPRMCANSSSLKIPVRYLAELHS--AQGLNFVHFAKSSFYTDITFGWIAQKLKTHLLAQ 259  
Db 253 L-----SFAAK-----QRELHSHVTNAMLVNVLQ-AIYVIDFF--W-----NE 288  
QY 260 TWQKKKQEL-----PSNCSLP-----Y 276  
Db 289 TWYLTIDICHDFHGWYLGWDCVWLPYLYTLOGLVLYVHPVQLSTPHAVGVLLGLVGY 348  
QY 277 HVYNI-----KSIGVTSKSYFSS--RODHSKMCVS--IKGSANRWTC 314  
Db 349 YIFRVANHQKDLFRRTDGRCLIWGRKPKVTECSVTSADGQRHSHKLLVSGFWGVARHFN 408  
QY 315 IGIDNRSLSHQLR-GGGFICTKHYIQAF-----HKLYLRG 351  
Db 409 VGDLMSLAYCLACGGHLLPYFYIYMAILTHRCLRDREHRCASKYG 456

RESULT 14  
US-09-817-774-26  
; Sequence 26, Application US/09817774  
; Patent No. US20020120111A1  
; GENERAL INFORMATION:  
; APPLICANT: CHOE, Sunghwa  
; APPLICANT: FELDMANN A, Kenneth  
; TITLE OF INVENTION: Dwf5 MUTANTS  
; FILE REFERENCE: 2225-0020 / 91020.002  
; CURRENT APPLICATION NUMBER: US/09/817,774  
; CURRENT FILING DATE: 2001-03-26  
; PRIOR FILING DATE: 2001-03-26  
; PRIOR APPLICATION NUMBER: 60/192,202  
; PRIOR FILING DATE: 2000-03-27  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 26  
; LENGTH: 475  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: S7R-HUMAN  
US-09-817-774-26

Query Match 5.3%; Score 101.5; DB 9; Length 475;  
Best Local Similarity 21.4%; Pred. No. 0.29;  
Matches 100; Conservative 48; Mismatches 147; Indels 173; Gaps 23;  
QY 6 LRTVLSLLFFA----LSGVLTGPISCRNEYGEAVDMF-----IPYKLPKRTSKASE 53  
Db 40 LASVIFLLFPFIVYFIMACDQVSCALT-GPVVDIVTGHARLSDIWAKTPPTRKAAQ 98  
QY 54 BAGL-----QYLYLDSRTQWNKSL-----YLNSTRSALGRTHQ 89  
Db 99 LYTLLVTFQVLLYTSLPDFCHKFLPGYVGGIOGAVTPAGVNVKQYINGLQAWLLTHLLW 158  
QY 90 LYDTHNSTNDTAYLIYNDGVPGSVNSYRQYGHAKGLLVNWKRTQGFWLIHSPKPPVHG 149  
Db 159 FANAHLWSFSTIIFDWNIP-----LLWCAN-----ILGAVSTFANVKGY 200  
QY 150 EYPTSGRRYGTGICITFGYSOFEEIDFQLLVLPQNIYSCF-----IPSTFWHKLIIY 201  
Db 201 FFPTSARDCKFTG---NFFYNMVGIEF-----NPRIGKWFDFKLFNGREGIVAWTLIN 252  
QY 202 MPRMCANSSSLKIPVRYLAELHS--AQGLNFVHFAKSSFYTDITFGWIAQKLKTHLLAQ 259  
Db 253 L-----SFAAK-----QRELHSHVTNAMLVNVLQ-AIYVIDFF--W-----NE 288  
QY 260 TWQKKKQEL-----PSNCSLP-----Y 276  
Db 289 TWYLTIDICHDFHGWYLGWDCVWLPYLYTLOGLVLYVHPVQLSTPHAVGVLLGLVGY 348  
QY 277 HVYNI-----KSIGVTSKSYFSS--RODHSKMCVS--IKGSANRWTC 314

```

Db      349 YIPRVANHQKDLFRRTDGRCLIWGRKPKVTECSVTGADGQRHSHKLLVSGFWGVARHFNY 408
QY      315 IGDNRSLHQAQR-GGGFICTKNHYIQAF-----HKLYLRYG 351
Db      409 VGDLMSLAYCLACGGGHLPLPYFIYIMAILLTHRCLRDDEHRCASKYG 456

```

```

RESULT 15
US-10-429-160-32
; Sequence 32, Application US/10429160
; Publication No. US20040023276A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Teresa R
; APPLICANT: Mao, Mao
; APPLICANT: Linsley, Peter S
; APPLICANT: Lund, Lund
; TITLE OF INVENTION: LXR Ligand Induced Genes and Proteins
; FILE REFERENCE: RS0200
; CURRENT APPLICATION NUMBER: US/10/429,160
; PRIOR FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 60/377,714
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-429-160-32

```

```

Query Match      5.3%; Score 101.5; DB 16; Length 475;
Best Local Similarity 21.4%; Pred. No. 0.29;
Matches 100; Conservative 48; Mismatches 147; Indels 173; Gaps 23;

```

```

QY      6 IRTVLSLLFFA----LSGVLTGPEISCRNEYGEAVDF-----IFYKLPRKTSKASE 53
Db      40 LASVIFLLFAPFIVYFIMACDQXSALT-GPVVDIVTGCHARLSDIWAKTPPIITRKAQ 98
QY      54 EAGL-----QVLYLDSTQWNSL-----YLINSTRSALGRTLOH 89
Db      99 LYTLMWTFQVLYTSLPDFCHFLPGVYGGIPEGAVTPAGVNVKYQINGLQAWLLTHLEW 158
QY      90 LYDTHNSTNDTAYLYNDGVPGSVNSROYGHAKGLLVNWRQTGGFWLIHSVPKFPFVHGY 149
Db      159 FANAHLLSWFSPTIIFDNWIP-----LLWCAN--ILGYAVSTFAMVKGY 200
QY      150 EYPTSGRRYQGTGICITFGYSOFERIDFOLLVQENIYSCF-----IPSTFHWKLIY 201
Db      201 FFPTSARDCKFTG---NPFYNTMMGIEF---NPRIGKWFDFKLFNGRPGIVAWTLIN 252
QY      202 MPRMCANSSSLKIPVRYLAELHS--AQGLNFVHPAKSPYTDITGTWIAQKLKTHLLAQ 259
Db      253 L-----SFAK-----QRELHSHVTNAMVLNVLQ-ALYVIDFP--W-----NE 288
QY      260 TWQKKQKQEL-----PNSCSLP-----Y 276
Db      289 TWYLTIDICHDFGWLGWDCVWLPLYTYLQGLYLVVHPVQLSTPHAVGVLLGLVGY 348
QY      277 HNYNI-----KSGVTSKSYFSS--RQHSKWCVS--IKGSANRWTC 314
Db      349 YIPRVANHQKDLFRRTDGRCLIWGRKPKVTECSVTGADGQRHSHKLLVSGFWGVARHFNY 408
QY      315 IGDNRSLHQAQR-GGGFICTKNHYIQAF-----HKLYLRYG 351
Db      409 VGDLMSLAYCLACGGGHLPLPYFIYIMAILLTHRCLRDDEHRCASKYG 456

```

```

Search completed: October 14, 2004, 00:45:57
Job time : 68.2152 secs

```

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2004, 00:08:06 ; Search time 28.8776 seconds  
(without alignments)  
1179.175 Million cell updates/sec

Title: US-10-790-589-2

Perfect score: 1927

Sequence: 1 MTAKPLRTVLSLFFALSGV.....KNHYIQAFHKLRYGVGCK 354

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78.\*

1: piri.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	1373	71.3	356	2	deoxyribonuclease
2	604.5	31.4	364	2	deoxyribonuclease
3	596.5	31.0	360	2	deoxyribonuclease
4	557.5	28.9	348	2	hypothetical prote
5	535	27.8	375	2	hypothetical prote
6	383	19.9	516	2	hypothetical prote
7	311	16.1	238	2	F098.2 protein -
8	106	5.5	622	2	fructanase - Bacte
9	105	5.4	1445	1	protein-tyrosine-p
10	104	5.4	344	2	secreted glycoprot
11	97	5.0	668	2	hypothetical prote
12	94	4.9	1385	2	parasporal crystal
13	94	4.9	2165	1	genome polyprotein
14	93.5	4.9	583	1	asparagine synthas
15	91.5	4.7	774	2	autolysin (amidase
16	91	4.7	1435	2	S54697
17	90	4.7	1276	2	T14271
18	89.5	4.6	2825	2	T18526
19	89	4.6	760	2	A99233
20	88.5	4.6	802	2	T21464
21	88.5	4.6	1309	2	F96509
22	88.5	4.6	1435	2	C90596
23	88	4.6	3388	1	GNWVDP
24	87.5	4.5	175	2	S73337
25	87.5	4.5	471	2	JC1403
26	87.5	4.5	912	2	T02892
27	87	4.5	900	2	C64232
28	87	4.5	2109	1	A46309
29	86.5	4.5	1198	2	S51434

30 86.5 4.5 1386 2 S73401  
31 85 4.4 867 1 GNLJSA  
32 85 4.4 1087 2 T30330  
33 85 4.4 1442 1 E48148  
34 84.5 4.4 282 2 F71806  
35 84.5 4.4 541 2 S57658  
36 84.5 4.4 566 2 T33042  
37 84 4.4 675 2 S60612  
38 84 4.4 5376 2 T42215  
39 83.5 4.3 394 2 JS0600  
40 83.5 4.3 524 2 G64243  
41 83.5 4.3 1064 1 S57450  
42 83.5 4.3 1409 1 OFFPCP  
43 83.5 4.3 1436 2 T14895  
44 83.5 4.3 1639 2 T50119  
45 83 4.3 415 2 B69875

## ALIGNMENTS

### RESULT 1

JC7131  
deoxyribonuclease I (EC 3.1.21.1) - rat (strain Wistar)  
N:Alternate names: deoxyribonuclease II-like acid deoxyribonuclease  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 09-Dec-2002  
C:Accession: JC7131  
R;Tanuma, S.; Shiohawa, D.  
Biochem. Biophys. Res. Commun. 265, 395-399, 1999  
A;Title: Cloning of a cDNA encoding a rat DNase II-like acid DNase.  
A;Reference number: JC7131; MUID:20025354; PMID:10558878  
A;Accession: JC7131  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-356 <TAN>  
A;Cross-references: GB:AF178974; NID:G6470130; PID:G6470131  
C;Superfamily: deoxyribonuclease II  
C;Keywords: hydrolase

Query Match 71.3%; Score 1373; DB 2; Length 356;  
Best Local Similarity 71.0%; Pred. No. 1.1e-109;  
Matches 252; Conservative 34; Mismatches 67; Indels 2; Gaps 1;  
QY 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEVGEAVDFYFKLPKRTSKASEAGLQYL 60  
Db 1 MTAKPLKAALPLLFVALSGVLGTPFVISCINEDGKAVDFYFKLPKRTSGGGTGMGLDYL 60  
QY 61 YLDSTQTNKSLYLINSTRSALGRTHLYDTHNSTNDTAVLIYNDVPGSVNYGRQYG 120  
Db 61 YLDSTMTWKSHELLINSSRSLGRTLEQLYEAHNAKNDTAVLIYNDVAPSVNYSGNYG 120  
QY 121 HAKGLLVNRTQGFVLIHSVPKPPV--HGVEYPTSGRRYGOTGICITFGYQFHEIDFQ 178  
Db 121 HAKGLLVNRTQGFVLIHSVPKPPV--HGVEYPTSGRRYGOTGICITFGYQFHEIDFQ 178  
QY 179 LLVQNIYSCFIPSTFTHKLIYMPKNCANSSSLKIPRYLAELHSAQGLNFVHFAKSSF 238  
Db 181 LLVQNIYSCFIPSTFTHKLIYMPKNCANSSSLKIPRYLAELHSAQGLNFVHFAKSSF 238  
QY 239 YDDIFTGWIQAOKLTHLLAQTWQKKQBLPNCSLPHVYVNIKSGTVTSKSYFSRQDH 298  
Db 241 YDDIFTGWIQAOKLTHLLAQTWQKKQBLPNCSLPHVYVNIKSGTVTSKSYFSRQDH 298  
QY 299 SKWCYSIKGSANRWTICIGDLNRSIHOALRGGGFICTKHYIYQAFHKLRYGVGFC 353  
Db 301 SKWCYSTKDSQARWTICIGDLNRSIHOALRGGGFICTKHYIYQAFHKLRYGVGFC 355

### RESULT 2

JF0205  
deoxyribonuclease II (EC 3.1.22.1) - pig  
C:Species: Sus scrofa domestica (domestic pig)

C;Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change 09-Dec-2002  
C;Accession: JE0205  
R;Shiokawa, D.; Tanuma, S.  
Biochem. Biophys. Res. Commun. 247, 864-869, 1998  
A;Title: Cloning of cDNAs encoding porcine and human DNase II.  
A;Reference number: JE0205; MUID:98321218; PMID:9647784  
A;Accession: JE0205  
A;Molecule type: mRNA  
A;Residues: 1-364 <SHI>  
A;Cross-references: GB:AF060221; NID:g3309153; PIDN:AAAC39263.1; PID:g3309153  
C;Comment: This enzyme catalyzes the hydrolysis of DNA into 3'-phosphoryl oligonucleotides  
C;Superfamily: deoxyribonuclease II  
C;Keywords: hydrolase

Query Match 31.4%; Score 604.5; DB 2; Length 364;  
Best Local Similarity 37.9%; Pred. No. 5.9e-44;  
Matches 136; Conservative 60; Mismatches 144; Indels 19; Gaps 9;

QY 6 LRTVSLIFPALSGV-LGTPEISCRNEYGEAVDFWFIYKLPKRTSKAS-BEAGLQYLVD 63  
DB 1 MATLSPLLLAALLWVPGT--LTCYDGGQGVDFWVYKLPKPAHSPGDVAQSLRYKILD 58  
QY 64 STROTWNKSLYLINSTSALGRITLQHLVDTHNSTNDTAYLIYNDGVP---GSVNYSRQY 120  
DB 59 EESGWRDAGSINSITCALGRSLPLY---RNTSQAFLYNDQPPKYRGS-QHSSNRG 114  
QY 121 HAKGLLVNRTQGFWLHSPKPPVH---CYEPTSGRRYGTGICITTCYSGQPEIDF 177  
DB 115 HTKGVLDDGGGFWLHSPVNFPPSSAAAYSPWPPSARTYGTQLICVSPFLTQFLISR 174  
QY 178 QLLVLQPNYISCFIPSTPHWKLIIYPRMCANSSSLKIPVRYLAELHSAQGLNFVHFVAKSS 237  
DB 175 QLTTPYPMYDYKLEGDFARFPYILEEVKGGHVLQEPWNSSVILTSCAGASFSQFACG 234  
QY 238 FYTDIDFTGMAQLKTHLLAQTWQKKQELPNSCLPIYHYNIKSIGVTSKS---YFSSR 295  
DB 235 NFGDDLISGWLALAGSNLQVQWORSAGILPNSCSGVQHVLDVQTQIAFPAGPAGFNAT 294  
QY 296 QDHSHKWCVSIKGSANRWTCIGDLNLSLHQAALRGGGFICTKHHYIQAAPHKLYRYGFC 354  
DB 295 EDHSHKWCVA---PERPMTCVGMNRNKEERHGGGTLCQAOLPALWKAFOPLVKNYQPC 350

RESULT 3  
JE0206  
deoxyribonuclease II (EC 3.1.22.1) - human  
C;Species: Homo sapiens (man)  
C;Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change 09-Dec-2002  
C;Accession: JE0206  
R;Shiokawa, D.; Tanuma, S.  
Biochem. Biophys. Res. Commun. 247, 864-869, 1998  
A;Title: Cloning of cDNAs encoding porcine and human DNase II.  
A;Reference number: JE0205; MUID:98321218; PMID:9647784  
A;Accession: JE0206  
A;Molecule type: mRNA  
A;Residues: 1-360 <SHI>  
A;Cross-references: GB:AF060222; NID:g3309154; PIDN:AAAC39852.1; PID:g3309155  
C;Comment: This enzyme catalyzes the hydrolysis of DNA into 3'-phosphoryl oligonucleotides  
C;Superfamily: deoxyribonuclease II  
C;Keywords: hydrolase

Query Match 31.0%; Score 596.5; DB 2; Length 360;  
Best Local Similarity 38.0%; Pred. No. 2.8e-43;  
Matches 132; Conservative 56; Mismatches 144; Indels 15; Gaps 7;

QY 17 LSGVLGTP--ETSCREYGEAVDFWFIYKLPK-RTSKASEAGLQYLVDSTROTWNKSL 73  
DB 6 LAALLCVPAGALTCTGDSGQVDFWVYKLPALRGSGEAAQGLQYKYLDESSGGWRDGR 65  
QY 74 YLINSTSALGRITLQHLVDTHNSTNDTAYLIYND--GVPGSVNYSRQYHAKGLLVNRT 131  
DB 66 ALINSPGAVGRSLQPLY--RSNTSQAFLYNDQPPQPSKAQDSSMRGHTKGVLLDHD 123  
QY 132 QGFWLHSPKPP---VHGVEYPTSGRRYGTGIC--ITTCYSQFEIDFOLLVLQPN 186  
DB 124 GGFVLHSPVNPFPPPASSAAYSWPHSACTYGTLLCKQLTYTY-----PWV 169  
QY 187 YSCFTPSFTHWKLIIYPRMCANSSSLKIPVRYLAELHSAQGLNFVHFVAKSSFYTDIDFTG 246  
DB 170 YNQLGEGFAQFPDLENVKGHVHVSQEPWNSITLTQAGAVFQSFQAKFSFGDDLXSG 229  
QY 247 WIAQKLKTHLLAQTWQKKQELPNSCLPIYHYNIKSIGVTSKS---YFSSRODHSKWCVS 304  
DB 230 WLAAALGTNLQVQFWHKTGVLIPNSCSDIWOVLNVNQIAFPAGPAGFSNSTDHSHKWCVS 289  
QY 305 TKGSSANRWTCIGDLNLSLHQAALRGGGFICTKHHYIQAAPHKLYRYGFC 353  
DB 290 PKGP---WTCVGMNRNKEERHGGGTLCQAOLPALWKAFOPLVKNYQPC 335

RESULT 5  
T19038  
hypothetical protein C07B5.5 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans

66 ALINSPGAVGRSLQPLY--RSNTSQAFLYNDQPPQPSKAQDSSMRGHTKGVLLDHD 123  
QY 132 QGFWLHSPKPP---VHGVEYPTSGRRYGTGICITTCYSQFEIDFOLLVLQPN 186  
DB 124 GGFVLHSPVNPFPPPASSAAYSWPHSACTYGTLLCVSFFFAQFSKMGKQLTYTYPWVYN 183  
QY 189 CFIPSTFTHWKLIIYPRMCANSSSLKIPVRYLAELHSAQGLNFVHFVAKSSFYTDIDFTGWI 248  
DB 184 YQLEGIFAQFPDLENVKGHVHVSQEPWNSITLTQAGAVFQSFQAKFSFGDDLXSGWL 243  
QY 249 AQKLKTHLLAQTWQKKQELPNSCLPIYHYNIKSIGVTSKS---YFSSRODHSKWCVS 306  
DB 244 AAALGTNLQVQFWHKTGVLIPNSCSDIWOVLNVNQIAFPAGPAGFSNSTDHSHKWCVS 303  
QY 307 GSANRWTCIGDLNLSLHQAALRGGGFICTKHHYIQAAPHKLYRYGFC 353  
DB 304 GP---WTCVGMNRNKEERHGGGTLCQAOLPALWKAFOPLVKNYQPC 347

RESULT 4  
T45071  
hypothetical protein R31240.2 [imported] - human  
C;Species: Homo sapiens (man)  
C;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Dec-2002  
C;Accession: T45071  
R;Lamerdin, J.; McCreedy, P.; Stillwagen, S.; Ramirez, M.; Carrano, A.  
submitted to the EMBL Data Library, November 1996  
A;Description: Characterization by genomic sequence analysis of a gene-rich 111 kb region  
A;Reference number: Z22906  
A;Accession: T45071  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-348 <LAM>  
A;Cross-references: EMBL:AD000092; PIDN:AA051172.1  
A;Experimental source: cell line 5HL2-B; fibroblast  
C;Genetics:  
A;Map position: 19p13.2  
A;Introns: 29/2; 89/3; 116/1; 159/1; 225/1  
C;Superfamily: deoxyribonuclease II

Query Match 28.9%; Score 557.5; DB 2; Length 348;  
Best Local Similarity 37.0%; Pred. No. 5.8e-40;  
Matches 129; Conservative 52; Mismatches 137; Indels 31; Gaps 9;

QY 17 LSGVLGTP--ETSCREYGEAVDFWFIYKLPK-RTSKASEAGLQYLVDSTROTWNKSL 73  
DB 6 LAALLCVPAGALTCTGDSGQVDFWVYKLPALRGSGEAAQGLQYKYLDESSGGWRDGR 65  
QY 74 YLINSTSALGRITLQHLVDTHNSTNDTAYLIYND--GVPGSVNYSRQYHAKGLLVNRT 131  
DB 66 ALINSPGAVGRSLQPLY--RSNTSQAFLYNDQPPQPSKAQDSSMRGHTKGVLLDHD 123  
QY 132 QGFWLHSPKPP---VHGVEYPTSGRRYGTGIC--ITTCYSQFEIDFOLLVLQPN 186  
DB 124 GGFVLHSPVNPFPPPASSAAYSWPHSACTYGTLLCKQLTYTY-----PWV 169  
QY 187 YSCFTPSFTHWKLIIYPRMCANSSSLKIPVRYLAELHSAQGLNFVHFVAKSSFYTDIDFTG 246  
DB 170 YNQLGEGFAQFPDLENVKGHVHVSQEPWNSITLTQAGAVFQSFQAKFSFGDDLXSG 229  
QY 247 WIAQKLKTHLLAQTWQKKQELPNSCLPIYHYNIKSIGVTSKS---YFSSRODHSKWCVS 304  
DB 230 WLAAALGTNLQVQFWHKTGVLIPNSCSDIWOVLNVNQIAFPAGPAGFSNSTDHSHKWCVS 289  
QY 305 TKGSSANRWTCIGDLNLSLHQAALRGGGFICTKHHYIQAAPHKLYRYGFC 353  
DB 290 PKGP---WTCVGMNRNKEERHGGGTLCQAOLPALWKAFOPLVKNYQPC 335

RESULT 5  
T19038  
hypothetical protein C07B5.5 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans

```

Qy 96 STNDTAYLLYNDGVPGSVNYSROYCHAKGLLLVWNRTOGFWLHHSVPKFPVPHGYEYPTSG 155
Db 280 NMTNTFSYMYNDSDWPDSTIWSNSGSHAGVTVDFDQYTGFWMIHSIPKFPKDMFRFESNA 339
Qy 156 RRYGQTGICITFGYSQFEEDIFOLLVLQPNYISCFIPSTFHWKLIYMPRMCANSSSLKIP 215
Db 340 HYTGQMGICISYNTVSLATTAQLFYNTFTYQFNLPQSFANQFPVLSQLKNKEYNKSPP 399
Qy 216 VRYLAEHLSAQGLNFVHFPAKSSPYTDIDFTGMTAQKLKTHLLAQTWOKK---KQELPSNC 272
Db 400 LTTSTKVLKSLGGQHFHFPAKTSWGDYLSDFVGPTLKSSIKVETWNHSGDEYNLPSCV 459
Qy 273 SLPHYVNYIKSIGVTSKSY---PSSRODSKWCVSIKGSANR-----WTGIDLNRLS 322
Db 460 D-PNHVQSTWSAKYIRLPDAIDYSSYEDHSKFVYVAYSESSKPPPIPVYVCGIDINQV 515

RESULT 7
S44793
F09G8.2 protein - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 14-Sep-1994 #sequence_revision 12-May-1995 #text_change 23-Mar-2001
C/Accession: S44793
R:Anderson, K.
submitted to the EMBL Data Library, February 1993
A/Description: Sequence of the C. elegans cosmid F09G8.
A/Reference number: S44776
A/Accession: S44793
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-238 <END>
A/Cross-references: EMBL:L11247; NID:gl56280; PID:gl56283
C/Genetics:
A/Introns: 32/2; 171/3

Query Match 16.1%; Score 311; DB 2; Length 238;
Best Local Similarity 29.5%; Pred. No. 4.3e-19;
Matches 76; Conservative 44; Mismatches 98; Indels 40; Gaps 10;

Qy 12 LLFPALSGVLGTPEISCRNEYGEADVDFYFKLPKRTSKASEAGLQLYLDSTQTNWK 71
Db 6 VLIFSIVFTTGGKIQCKNMRGKSDVFWVYKLPKLSGAGT--SGKEFYVFAESSDWTR 63
Qy 72 SLYLINSTSALGRTHLOHLYDHNSTNDTAYLIYNDGVPGSVNYSROY-CHAKGLLVWNR 130
Db 64 G-NDINDPNVAVGATQSVYSADKSN--FWFMYSDDDP--IKSADSYRGHAKGVSLFDS 118
Qy 131 TQGFLLHSPKPPVPHGYEYPTSGRYGOTGICITFGYSQFEEDIFOLLVLQPNYISCF 190
Db 119 TTGFLLHSPVFPFPKPSYNTAEKYQSFCAEVSQFCAEVSQFCAEVSQFCAEVSQFCAE 163
Qy 191 IPSTFHWKLIYMPRMCANSSSLKIPVRYLAELHLSAQGLNFVHFPAKSSPYTDIDFTGMT 250
Db 164 ---AEHWKIYQ----ATPYIINPEKYATRPFLKQVE---AKQSL-PRSATQFWISK 210
Qy 251 KLKT-----HLLAQTWOKK 264
Db 211 PIKTQVQVTIMAYAKHK 228

RESULT 8
A36915
fructanase - Bacteroides fragilis
C/Species: Bacteroides fragilis
C/Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 17-Mar-2000
C/Accession: A36915
R:Blatch, G.L.; Woods, D.R.
J. Bacteriol. 175, 3058-3066, 1993
A/Title: Molecular characterization of a fructanase produced by Bacteroides fragilis BF-1
A/Reference number: A36915; NUID:93259952; PMID:8491724
A/Contents: BF-1
A/Accession: A36915

```

A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-622 <BLA>  
A;Cross-references: GB:M83774; NID:g143370; PIDN:AAA22924.1; PID:g143372  
A;Note: sequence extracted from NCBI backbone (NCBIN:131947, NCBIP:131949)  
C;Superfamily: beta-fructofuranosidase

Query Match 5.5%; Score 106; DB 2; Length 622;  
Best Local Similarity 21.8%; Pred. No. 0.5;  
Matches 79; Conservative 34; Mismatches 142; Indels 108; Gaps 19;

QY 34 EAVDVFIFVYKLPRTSKASEAGLOLYLDSTQRTWKSILYL---INSTRSALGRITLOHL 90  
DB 84 EKVDYFVPALP-----KGEKAVAVRIHLEPKALCW-KELKLSDTFTTNDQYRPLYXH 138  
QY 91 YDTHNSTNDYALILYNDGVPGSVNYRQYGHAKGLLVNRTQFVLIHSPKPPVHYGE 150  
DB 139 TPLYGWNMDANGILVYKDGHYHLYFYQYNPNFYSGMWNHMG-----HSVSK-DLVH-WE 188  
QY 151 Y-----PTSGRRYQGTGI---CITFGYSQFEIDFQLLVLPN-IYSCFIPSTFWKL 199  
DB 189 HLEPALARDTLGHIFGSSVVDANTAGYAGAVFYTSASDKNGQIQCMAYSTNGRT 248  
QY 200 IYPMRCANSSSLKIPVRYLAELHSAQGLNFVFAKSSFTYDDIFTGWI----- 248  
DB 249 F-----TKYEKNV-----LTPDGLKDFRDPKVFVAPD--QKWMVVSADKEMRF 293  
QY 249 --AQKLTLLAQTW-----QKKQELPNSCLPVLHYVNIKIGVTSKYSFYSSRODHSK 300  
DB 294 YSSENLEKWTYMSGEGYGVQSPQCEPDVPEL-----VDGNPDHKK 337  
QY 301 WCYSIK-----GSANRWTCIGDNLRLHQLRGGGFICT-----KHHYIY 340  
DB 338 WALIVNPNPCYGGSGATQY-FIGDFD-----GEKFCVCKNPETVKGLDWGKHAT 388  
QY 341 QAF 343  
DB 389 VCF 391

RESULT 9  
A48148  
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type gamma - human  
C;Species: Homo sapiens (man)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C;Accession: A48148; C36065; S12051  
R;Barnea, G.; Silvennoinen, O.; Shaanan, B.; Honegger, A.M.; Canoll, P.D.; D'Eustachio, Mol. Cell. Biol. 13, 1497-1506, 1993  
A;Title: Identification of a carbonic anhydrase-like domain in the extracellular region  
A;Reference number: A48148; MUID:93180796; PMID:8382771  
A;Accession: A48148  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-1445 <BAR>  
R;Kaplan, R.; Morse, B.; Huebner, K.; Croce, C.; Howk, R.; Ravera, M.; Ricca, G.; Jaye, Proc. Natl. Acad. Sci. U.S.A. 87, 7000-7004, 1990  
A;Title: Cloning of three human tyrosine phosphatases reveals a multigene family of receptors  
A;Reference number: A36065; MUID:90384936; PMID:2169617  
A;Accession: C36065  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 874-1118,1175-1409 <KAP>  
R;Kueger, N.X.; Streuli, M.; Saito, H. EMBO J. 9, 3241-3252, 1990  
A;Title: Structural diversity and evolution of human receptor-like protein tyrosine phosphatases  
A;Reference number: S12049; MUID:91006018; PMID:2170109  
A;Accession: S12051  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 836-1406, M, 1408-1445 <KRU>  
A;Cross-references: GB:X54132; NID:g35793; PIDN:CAA38067.1; PID:g35794  
C;Genetics:

A;Gene: GDB:PTPRG; D3S1249  
A;Cross-references: GDB:127351; OMIM:176886  
A;Map position: 3p14.2-3p14.2  
C;Superfamily: protein-tyrosine-phosphatase, receptor type gamma; carbonic anhydrase homone-phosphatase homology  
C;Keywords: phosphoprotein; phosphoric monoester hydrolase; receptor; transmembrane protein-tyrosine-phosphatase homology <CAH>  
F;60-321/Domain: carbonic anhydrase homology <CAH>  
F;346-434/Domain: fibronectin type III repeat homology <3FN>  
F;742-758/Domain: transmembrane #status predicted <TM>  
F;874-1108/Domain: protein-tyrosine-phosphatase homology <PTP>  
F;1175-1399/Domain: protein-tyrosine-phosphatase homology <PTP2>  
F;1060/Active site: Cys (phosphocysteine intermediate) #status predicted  
F;1066/Binding site: substrate phosphate (Arg) #status predicted

Query Match 5.4%; Score 105; DB 1; Length 1445;  
Best Local Similarity 18.4%; Pred. No. 1.8;  
Matches 60; Conservative 51; Mismatches 89; Indels 126; Gaps 17;

QY 14 FFALSGVLG-----TPEISCRNEYGAVDWFIYFKLPKRTSKASEAGLOLYLDSTRTQ 68  
DB 59 YWAYSGAYGPEHWVTSSVSGSRHQSPIDLDQY---ARVGEYQELQLDGFNDSNKT 115  
QY 69 WNKSLYLINSTRSALORTLQHLTYDTHNSTNDYALILYNDGVPGSVNYSR---QYGHAKGL 125  
DB 116 WKNV-----TGKTVAILLKDD-----YFVSGAGLPGRFKAKEVFEFHGHSNGS 158  
QY 126 LVNWRQTGFWLHVSVPKFPVHGVEYPTSGRRYQGTGICITFGYSQFEEDFOLLVLPN 185  
DB 159 A-----GSEHSINGRR-----PVEMQIFPNPD 182  
QY 186 IYSCFIPSTFWKL-----YPMRCANSSSLKIPVRYLAELHSAQGLNFVFAKSS 237  
DB 183 DFDSPQTALSEWELIIGAMAFQVSPR---DNSALD-PI-----IHLKGV--VHEKET 231  
QY 238 F-----YTDDIFTGWIQAOKLKLHLLAQTWQKKQELPNSCSPIYH- 277  
DB 232 FLDPFVRLDLPASLGSYYRTGSLTPPCSEIVE-----WIVFRPVP-----ISYHQ 280  
QY 278 ---VYNIKSIGVTSKYSFYSSRODHSK 300  
DB 281 LEAFYSI-----FTTEQQDHVK 297

RESULT 10  
A44164  
secreted glycoprotein, 43 K - trichina  
C;Species: Trichinella spiralis (trichina)  
C;Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 09-Sep-1997  
C;Accession: A44164; A60630; S27862  
R;Vassilatis, D.K.; Despomnier, D.; Misk, D.E.; Polvere, R.I.; Gold, A.M.; Van der Ploeg, J. Biol. Chem. 267, 18459-18465, 1992  
A;Title: Analysis of a 43-kDa glycoprotein from the intracellular parasitic nematode Trichinella spiralis  
A;Reference number: A44164; MUID:92406752; PMID:1382055  
A;Accession: A44164  
A;Molecule type: mRNA  
A;Residues: 1-344 <VAS>  
A;Cross-references: EMBL:M95499; NID:g162534; PID:g162535  
A;Note: sequence extracted from NCBI backbone (NCBIN:113310, NCBIP:113321)  
A;Note: part of this sequence was confirmed by protein sequencing  
R;Gold, A.M.; Despomnier, D.D.; Buck, S.W. Mol. Biochem. Parasitol. 41, 187-196, 1990  
A;Title: Partial characterization of two antigens secreted by L1 larvae of Trichinella spiralis  
A;Reference number: A60630; MUID:90377287; PMID:2398916  
A;Accession: A60630  
A;Molecule type: protein  
A;Residues: 23, X, 25-62, YGSP, 66-80, T, 115-116, X, 118-132; 249-278 <GOL>  
C;Keywords: glycoprotein

Query Match 5.4%; Score 104; DB 2; Length 344;  
Best Local Similarity 19.0%; Pred. No. 0.35;  
Matches 65; Conservative 60; Mismatches 119; Indels 98; Gaps 17;

QY 27 SCRNEYGAVDWFIYFK-----LPKRTSKASEAGLOLYLDSTRTWKNLSYLINSTRSAL 83



Db 26 TCKTATDD-TWELFKPVGLLKAKIISPANACWAN---DGANWNIDSGHALVQTLEAWM 81  
QY 84 GRTLOHLYTHNSTNTAVILYNDGVPGVNYSRQYGHAKGLLWNR--TQGFWLHISVP 141  
Db 82 GPIL-----DDMTALGYSNTPPKS-TITSTSSKGLMFGNETTDGFWLLHTFE 130  
QY 142 K-FPPVHGVEYPTSGRRYGTGICITFGYS-----QFEIDFQLLVQENIYSCFI 191  
Db 131 RAFFNSVANSWSKFTSEGHMALCISEDNVPLIVPALQIQSV----- 174  
QY 192 PSTFHWKLIYPMRC-----ANSSSL--KIPV-----RYLAELHSAQGLNFVHF 233  
Db 175 -----VIYFGQVSSEKATEFADLTSLIDGSLPTITPPLWNOQTITLNSA--LSTVVY 225  
QY 234 AKSSFYTDIDIFGWIAQKLTLLHQAQWQKKQELPSNLSPLHVNIKISGVTSKSYPS 293  
Db 226 SKTSSSRLEWYGSFLAKVMVNM--RIWAVTDNTLQTCGKIGFVKVVKSPVTTIDGTQ 283  
QY 294 SR-QDHSKWCV-----SIKGSANRWTCI 315  
Db 284 DRSKDKSQNAVIDDKPVFCFTNGYSTKQRTVAGSA---TCI 322

RESULT 11  
T18635  
hypothetical protein B0019.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T18635  
R:Kershaw, J.  
submitted to the EMBL Data Library, November 1997  
A:Reference number: Z19000  
A:Accession: T18635  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-668 <WIL>  
A:Cross-references: EMBL:AL008866; PIDN:CAAL5509.1; GSPDB:GN00019; CESP:B0019.1  
A:Experimental source: clone B0019  
C:Genetics:  
A:Gene: CESP:B0019.1  
A:Map position: 1  
A:Introns: 42/3; 103/1; 144/2; 195/3; 258/1; 304/1; 357/1; 404/2; 457/2; 474/1; 589/1; 6

Query Match 5.0%; Score 97; DB 2; Length 668;  
Best Local Similarity 20.0%; Pred. No. 3.2; Mismatches 98; Indels 136; Gaps 16;  
Matches 69; Conservative 42; Mismatches 98; Indels 136; Gaps 16;  
QY 5 PLRTVLSLLFFALSGVLGTPFEISCRNEYGEAVDWFIFKLPKRTSKASBEAGLQVLYLDS 64  
Db 109 PTDLTQLLSLQELN-VRTTQQLTCGNN-----TVFQQRKKQL-----S 147  
QY 65 FROTWNKSIY--LINSRS-----ALGRVLOHLYD----- 92  
Db 148 LQQQSTTLFTDLINSPELGLNTNTSVSAMSQOMDTADSDVNRNMQTFDFAPGEQVPE 207  
QY 93 -----THNSTNDTAVILYNDGVPGVNYSRQYGHAKGLI-----VWNRQGGFWLHIS 139  
Db 208 LQALFTCSQNAFA-----VEILRRGHQSLQAQGMNEVVRRLADGLLIYS 256  
QY 140 VKPFPVHGVEYPTSGRRYGTGICITFGYSQFEID-----FQLLVLQ--ENIYSC 189  
Db 257 QPAFAV---VQTSAGRRSGRQVIVAVPIPTLENIELVPAPEAPFQQLIYNGTGHAY 313  
QY 190 FIPSTFH--W-----KLIY-----MPCWANSLSK----- 213  
Db 314 YFTMSFORATWRLNGSRGKVIYTSATGPLVWLTTDTTPAASCDNSTSASSTLWGLAHP 373  
QY 214 ---IPVRYLAELHSAQGLNFVHFAKS-----FYTDDIFTGWI 248  
Db 374 YDVPETRRKLYQAIMYSLRFAFSPLDVSDVNFATDDLAKGTI 418

## RESULT 12

T18213  
parasporal crystal protein cry5Aa1 - Bacillus thuringiensis (fragment)  
C:Species: Bacillus thuringiensis  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 01-Dec-2000  
C:Accession: T18213  
R:Narva, K.E.; Payne, J.M.; Schwab, G.E.; Hickie, L.A.; Galasan, T.; Sick, A.J.  
submitted to the EMBL Data Library, July 1999  
A:Description: Novel Bacillus thuringiensis microbes active against nematodes, and genes  
A:Reference number: Z18829  
A:Accession: T18213  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1385 <NAR>  
A:Cross-references: EMBL:L07025; NID:g142868; PID:g142869; PIDN:AAA67694.1  
C:Genetics:  
A:Gene: CryVA (a)

Query Match 4.9%; Score 94; DB 2; Length 1385;  
Best Local Similarity 20.5%; Pred. No. 15; Mismatches 76; Conservative 48; Mismatches 134; Indels 112; Gaps 16;  
QY 31 EYGEAVDWFIFKLPKRTSKASEEAGLQVLYLDSR-----QTWNK----- 71  
Db 399 EVQDSVETRLYQLP---AVDPQAGPNVVSIDSNPILQINNDTWKTPPQGSAGWNTNL 454  
QY 72 -----SLYLINSTRSALGR-----TLQHYDTHNSTNDTAYLIYNDGVP--GSVNY 115  
Db 455 MRGSVGLSFLORDGTRLSAGMGGFADTIYSLPATH-----YLSYLYGTPTPYQTSNDY 507  
QY 116 SRQYGHAKGLLVNRTQGFWLHISVPK---FPVVGVEYPTSGRRYGTGICITFGYSQF 172  
Db 508 SGHVGLVGV-----STPQEAFLNIIG-----QDEQGNVSTMGF-PF 545  
QY 173 BEIDFQLLVLPNIYSCFIPSTFHWKLIYMPKANSLSKIPVRYLAELHS----- 224  
Db 546 EKASYGGTVKKEWLNAGANAKLSPOSGIGIPITNTSGEYQIRCRYASNDNTVNFVNT 605  
QY 225 -----AQGLNFVHFAKSSPYTDDIFTGWTIAQKLKTHLLAQ----- 259  
Db 606 GGANPIFQOINFASTVDNNTGVQANGVYVVKRSIATDTSFTETPAKTINVLHNTQSSD 665  
QY 260 TWQKKKQELPSNCSIPYH-VYNIKISGVTSKSYFSSRQDHSKWCYSIKGSANRWTCIGDL 318  
Db 666 VFDRIEFIPFSLPIYHGSYNTSS-GADDVLWSSSNMNY--YDIIVNGQANS----SSI 718  
QY 319 NRSLSQALRG 328  
Db 719 ASSMHLNKG 728

## RESULT 13

RRN2A2  
genome polyprotein - human respiratory syncytial virus (strain A2)  
N:Alternate names: polymerase L protein  
N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)  
C:Species: human respiratory syncytial virus  
A>Note: host Homo sapiens (man)  
C>Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 16-Jun-2000  
C:Accession: A40317; A28319; PS0048  
R:Stec, D.S.; Hill III, M.G.; Collins, P.L.  
Virology 183, 273-287, 1991  
A:Title: Sequence analysis of the polymerase L gene of human respiratory syncytial virus  
A:Reference number: A40317; MUID:9127488; PMID:2053282  
A:Accession: A40317  
A:Molecule type: mRNA  
A:Residues: 1-2165 <STE>  
A:Cross-references: GB:M75730; NID:g333955; PIDN:AAA47418.1; PID:g333956  
R:Collins, P.L.; Olmsted, R.A.; Spriggs, M.K.; Johnson, P.R.; Buckler-White, A.J.  
Proc. Natl. Acad. Sci. U.S.A. 84, 5134-5138, 1987  
A:Title: Gene overlap and site-specific attenuation of transcription of the viral polymerase  
A:Reference number: A28319; MUID:87260943; PMID:2440043  
A:Accession: A28319

A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-81 <COL>  
A:Cross-references: GB:M17245; NID:g333953; PIDN:AAA47417.1; PID:g333954  
R:Johnson, P.R.; Collins, P.L.  
J. Gen. Virol. 69, 2901-2906, 1988  
A:Title: The A and B subgroups of human respiratory syncytial virus: comparison of inter-  
A:Reference number: PS0048; MUID:89036169; PMID:3183631  
A:Accession: PS0048  
A:Molecule type: mRNA  
A:Residues: 1-18 <JOH>  
A:Cross-references: GB:D00397; NID:g222551; PID:g2160375  
A:Experimental source: strain 18537  
A:Note: this strain belongs to subgroup B  
C:Genetics:  
A:Gene: L  
C:Keywords: Atp; nucleotidyltransferase  
C:Keywords: Atp; nucleotidyltransferase

Query Match 4.8%; Score 94; DB 1; Length 2165;  
Best Local Similarity 21.5%; Pred. No. 27; Mismatches 92; Indels 86; Gaps 12;  
Matches 58; Conservative 34

QY 112 SVNRYGYCHAKGLLVNRRYQGFWLHSPKFPVHGYEPTS-----GRRYQGTGI-- 163  
DB 1333 SVNLYHR-----LTVSSRPECF--PASIPAYRTTN-YHFDTPSPINRLTEKYGDEDIDI 1383

QY 164 ----CITFGYSQFEIDFOLLVQPNYSFCIPFTFWKLIYMPRCANSSSLKIPVRYL 219  
DB 1384 VFQNCISFGLSLMSVVE-QFTNVCPN-----RIILIPK----- 1415

QY 220 AELHSAAGLVNHFPAKSSFYDDITFTGWIQAQKLKTHLLAQWQKKQELPNSCLPYHYV 279  
DB 1416 -----LNEIHLMKPPIFTGDV-----DIHLKLVQIQKHMFELDKISLTQV- 1457

QY 280 NIKSIGTYSYFSSRQDHKWCVSIKGS-----ANRWTCIGDL-----NRSLLH 323  
DB 1458 ---ELFSLNKLTKSGSHVNSNLIHLKSIDYFHTYILSTNLAGHWILLIQLMKDSKGIP 1514

QY 324 QALRGGFICTKNHYIYQAFHKLVRYGFC 353  
DB 1515 EKWGEGYITDHMPINLKVFFNAYKTYLLC 1544

RESULT 14  
A:JPWN2  
asparagine synthase (glutamine-hydrolyzing) (EC 6.3.5.4) [similarity] - garden pea  
N:Alternate names: asparagine synthetase (glutamine-hydrolyzing)  
C:Species: Pisum sativum (garden pea)  
C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 02-Aug-2002  
C:Accession: S11443  
R:Tsai, F.Y.; Coruzzi, G.M.  
EMBO J. 9, 323-332, 1990  
A:Title: Dark-induced and organ-specific expression of two asparagine synthetase genes i  
A:Reference number: S11443; MUID:90151604; PMID:1968003  
A:Accession: S11443  
A:Molecule type: DNA  
A:Residues: 1-583 <TSA>  
A:Cross-references: EMBL:X52180; NID:g20651; PIDN:CAA36430.1; PID:g20652  
C:Genetics:  
A:Gene: AS2  
C:Superfamily: asparagine synthase (glutamine-hydrolyzing)  
C:Keywords: asparagine biosynthesis; ligase  
P/2-583/Product: asparagine synthase (glutamine-hydrolyzing) #status predicted <MAT>  
P/2/Active site: Cys #status predicted

Query Match 4.9%; Score 93.5; DB 1; Length 583;  
Best Local Similarity 20.4%; Pred. No. 5.4;  
Matches 80; Conservative 45; Mismatches 130; Indels 137; Gaps 20;

QY 24 PEISCRNEYGAVDWFIYKLPKRTSKASBAGLOLYLDSTQ--TWKSLYLINSTRS 81  
DB 33 PEWSGLHQHGDG-----YLAQORLAIYDPASGDQPLEFNEPNPSIVTVNGEINYHEDLRK 86

QY 82 ALGR-----TLQHLTYDTH-----STNDTAYLIYNDGVPGSVNY 115  
DB 87 QLSNHFRTGSCDCAVIAHLYEEGEFVMDLGGIFSVPLDTRNSVIIVARDAI----- 140

QY 116 SROYGHAKGLLVNRRYQGFWLHSPK-----FPVHGYEYPTSG-RRYGGQTGIC 164  
DB 141 ----GVTSLYIGWGLDGSWISSEMGLNDCBHFCEFPHPGHLYSSKDSGFRW----- 190

QY 165 ITFGYSQFEIDFOLLVQPNYSFCIPFTF-----WKLIYMPRCAN----- 208  
DB 191 -----YNPSWYSEALPSAPYDPLALRHAFKAVVRLMTDVPFGVLLS 233

QY 209 ---SSSL--KIPVRYLAELHSAQ--GLNFVHFVAKSSFYDDITFTG-WIAQKLKT--HLLA 258  
DB 234 GGLDSSLVASITSRYLATTKAABQWGSKLHSCFVGLGSPDLKAGKEVADYLGTVHHEFT 293

QY 259 QTWQKKQKELPNSCLPYHV--YNIKSIGTYSYFSSRQDHS-----KWCVSIKGSANRW 313  
DB 294 FTVDGDIDATE--DVIYHVETVDVTSIRASTPMFLMSRKIKSLGVKRWISGEGS----- 345

QY 314 CIGDLNPSLHQAALRGGFICTKNHYIYQAFHK 345  
DB 346 -----DEIFGG-----YLY--PHK 357

RESULT 15  
AG1565  
autolysin (amidase) homolog lin1064 [imported] - Listeria innocua (strain Clip11262)  
C:Species: Listeria innocua  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AG1565  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,  
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Psihi, H.;  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mat  
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,  
A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AG1565  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-774 <GLA>  
A:Cross-references: GB:AL592022; PIDN:CAC96295.1; PID:g16413523; GSPDB:GN00178  
A:Experimental source: strain Clip11262  
C:Genetics:  
A:Gene: lin1064

Query Match 4.7%; Score 91.5; DB 2; Length 774;  
Best Local Similarity 20.6%; Pred. No. 12;  
Matches 68; Conservative 56; Mismatches 137; Indels 69; Gaps 18;

QY 1 MTAKPLRTVLSLLFPALSGVLGTEISCRNEYGAVDWFIYKLPKRTSKA-SHEAGI-QY 59  
DB 363 LTIYFSDSVISSKAVNLDGQITNPT-----NGGI-WTKAYKLEGTTSVAQATQYANQD 414

QY 60 LYLDSTRTQWNKSLYLINSTRSALG--RTLQHLTYDTHNSTNDTAYLIYNDGV--PGSVN 114  
DB 415 VKISEVETQGTYYNISIGRAIGWLDKNAITLYDQAE-----YKNAVTLDTGK 465

QY 115 YSROYGHAKGLLVNRRYQGFWLHSPK-----PPVHGYEYPTSGRRY 158  
DB 466 -----NVQGNAIWTEPYRTAGTKLVAQAEKYANKDVQIVREAKTPRGTYQYFSGGNVI 519

QY 159 GQTGICITFGYSQFEIDFOLLVQPNYSFCIPFTFWKLIYMP--RMCANSSSLK-I 214  
DB 520 GWLDV---KAPDMYDEITSNKAVNNEAVIE-NVSGNAVWTPAYKSVGKLIGPASNYDK 575

QY 215 PVRVLAELHSAQGLNVHFPAKSSFYDDITFTGWIQAQK-LKTHLLAQWQK-KKQELPNSC 272  
DB 576 TVKLTREAQTSRGTY-----EFSSNGKTVGWLDKKAFKTYSTIYDESIVRDVAVTVN 629

OY 273 S-----LPYHVYNIKSIGVTSKSYFSSRQ 296  
| | | | | : | | | | : : : :  
Db 630 SGHTVWTLPYQYGVKSVG--SASNYNNKQ 657

Search completed: October 14, 2004, 00:32:21  
Job time : 31.8776 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 22:11:29 ; Search time 71.1983 Seconds  
(without alignments)  
1404.834 Million cell updates/sec

Title: US-10-790-589-2

Perfect score: 1927

Sequence: 1 MTAFLRTVLSLFFALSGV.....KNHYIQAPKLYLRYGCK 354

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*

1: Geneseqp1980s.\*

2: Geneseqp1990s.\*

3: Geneseqp2000s.\*

4: Geneseqp2001s.\*

5: Geneseqp2002s.\*

6: Geneseqp2003as.\*

7: Geneseqp2003bs.\*

8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1927	100.0	354	4	AAB72416 Murine DN
2	1927	100.0	354	4	AAB72416 Murine DN
3	1293.5	67.1	357	4	AAB13014 Mouse deo
4	1293.5	67.1	357	5	AAB13015 Human deo
5	1286.5	66.8	361	4	ABG61821 Prostate
6	596.5	31.0	360	2	AAB72417 Human DNA
7	596.5	31.0	360	2	AAB72417 Human DNA
8	444	23.0	366	4	ABG63496 Drosophil
9	443	23.0	276	2	AAB72921 Bos tauru
10	236.5	12.3	192	6	ABG72082 Mouse end
11	101.5	5.3	475	2	AAY29333 Human sec
12	101.5	5.3	475	2	AAY29333 Human sec
13	101.5	5.3	475	2	AAB93573 Human Del
14	101.5	5.3	475	4	AAB93572 Human Del
15	101.5	5.3	475	5	AAB39059 Human sec
16	101.5	5.3	475	6	ABU89736 Human pol
17	101	5.2	481	6	ADA36462 Acinetoba
18	97	5.0	419	2	AAB06703 Medium ch
19	96	5.0	2165	5	ABG67251 Respirato
20	95	4.9	2499	6	ABM70221 Phototrab
21	94.5	4.9	2165	5	ABG67249 Respirato
22	94	4.9	1385	2	AAR29516 Br toxin
23	94	4.9	1385	2	AAR20066 B.thuring
24	94	4.9	1385	2	AAB59881 17a prote
25	94	4.9	1385	2	AAR29026 Bacillus

## ALIGNMENTS

RESULT 1

AAB72416

ID AAB72416 standard; protein; 354 AA.

XX AC AAB72416;

XX DT 03-MAY-2001 (first entry)

XX XX Murine DNase.

XX DE Murine DNase.

XX KW Murine; DNase; DLAD; deoxyribonuclease; endonuclease; cystic fibrosis;

XX KW infectious disease.

XX OS Mus musculus.

XX PN WO200112793-A1.

XX PD 22-FEB-2001.

XX PF 01-MAY-2000; 2000WO-JP002893.

XX PR 17-AUG-1999; 99JP-00230870.

XX PA (TANU/) TANUMA S.

XX PI Tanuma S, Shiokawa D;

XX DR WPI; 2001-218348/22.

XX PT N-PSDB; AAF60715.

XX PT Acidic deoxyribonuclease capable of divalent cation-independent cleavage of DNA under acidic even neutral pH and not inhibited by G-actin, useful in remedies for cystic fibrosis and for treatment of infectious diseases.

XX PS Claim 3; Page 48-49; 61pp; Japanese.

XX CC The present sequence is a murine deoxyribonuclease (DLAD), which is an endonuclease. DLAD is capable of divalent cation-independent cleavage of DNA under acidic conditions. This protein can be used as a substitute for DNase I in treating cystic fibrosis, and is useful in the prevention and treatment of infectious diseases

XX SQ Sequence 354 AA;

Query Match 100.0%; Score 1927; DB 4; Length 354;

Best Local Similarity 100.0%; Pred. No. 9.1e-191;

Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTAKPLRTVLSLLFFALSGVLGTPETSCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYL 60  
 Db 1 MTAKPLRTVLSLLFFALSGVLGTPETSCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYL 60  
 QY 61 YLDSTRQTWNKSLYLINSTRSALGRTLQHLHYDTHNSTNDTAYLIYNDGVPGSVNYSRQYG 120  
 Db 61 YLDSTRQTWNKSLYLINSTRSALGRTLQHLHYDTHNSTNDTAYLIYNDGVPGSVNYSRQYG 120  
 QY 121 HAKGLLVNRTQGFWLIHSVPKPPVHGYEYPTSGRRYQGTGICITFGYSQFEEDFOLL 180  
 Db 121 HAKGLLVNRTQGFWLIHSVPKPPVHGYEYPTSGRRYQGTGICITFGYSQFEEDFOLL 180  
 QY 181 VLOPNYSCFIPSTFHWKLIYMPRMCANSSSLKIPVRYLAELHSAOGLNFVHFAKSSFYT 240  
 Db 181 VLOPNYSCFIPSTFHWKLIYMPRMCANSSSLKIPVRYLAELHSAOGLNFVHFAKSSFYT 240  
 QY 241 DDIFTGWIAQKLTHLLAQTWKKQELPNCSLPYHVNIKSGIVTSKSYFSSRDHSHK 300  
 Db 241 DDIFTGWIAQKLTHLLAQTWKKQELPNCSLPYHVNIKSGIVTSKSYFSSRDHSHK 300  
 QY 301 WCVSIKGSANRWTCIGDLNRSLSHQALRGGGFICTKNHYIYQAFHKLYLRYGFCCK 354  
 Db 301 WCVSIKGSANRWTCIGDLNRSLSHQALRGGGFICTKNHYIYQAFHKLYLRYGFCCK 354

RESULT 2  
 AAE13014  
 ID AAE13014 standard; protein; 354 AA.  
 AC AAE13014;  
 XX  
 XX 28-JAN-2002 (first entry)  
 DT  
 DE Mouse deoxyribonuclease (DNase) II beta protein.  
 XX  
 XX Mouse; deoxyribonuclease; DNase II beta protein; mucous plug; lung;  
 KW cystic fibrosis; pulmonary.  
 XX  
 XX Mus sp.  
 XX  
 XX WO200175082-A1.  
 XX  
 XX 11-OCT-2001.  
 XX  
 XX 02-APR-2001; 2001WO-US010635.  
 XX  
 XX 03-APR-2000; 2000US-00541840.  
 XX  
 XX 19-MAY-2000; 2000US-00574942.  
 XX  
 XX (DART-) DARTMOUTH COLLEGE.  
 XX  
 XX Eastman AR, Krieser RJ;  
 XX  
 XX WPI; 2001-662972/76.  
 XX  
 XX N-PSDB; AAD21288.  
 XX  
 XX New cDNA encoding a deoxyribonuclease II beta enzyme useful for degrading  
 PT DNA present in the mucous plugs in the lungs of cystic fibrosis patients.  
 XX  
 XX Claim 5; Page 14-15; 21pp; English.  
 XX  
 XX The invention relates to deoxyribonuclease (DNase) II beta proteins and  
 CC their corresponding cDNAs. The DNase II beta may be useful to digest DNA  
 CC in the mucous plugs in lungs of cystic fibrosis patients and so reduce  
 CC their viscosity. The present sequence is mouse DNase II beta protein  
 CC their viscosity. The present sequence is mouse DNase II beta protein  
 XX  
 XX Sequence 354 AA;  
 SQ  
 Query Match 100.0%; Score 1927; DB 4; Length 354;  
 Best Local Similarity 100.0%; Pred. No. 9.1e-191;  
 Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTAKPLRTVLSLLFFALSGVLGTPETSCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYL 60

Db 1 MTAKPLRTVLSLLFFALSGVLGTPETSCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYL 60  
 QY 61 YLDSTRQTWNKSLYLINSTRSALGRTLQHLHYDTHNSTNDTAYLIYNDGVPGSVNYSRQYG 120  
 Db 61 YLDSTRQTWNKSLYLINSTRSALGRTLQHLHYDTHNSTNDTAYLIYNDGVPGSVNYSRQYG 120  
 QY 121 HAKGLLVNRTQGFWLIHSVPKPPVHGYEYPTSGRRYQGTGICITFGYSQFEEDFOLL 180  
 Db 121 HAKGLLVNRTQGFWLIHSVPKPPVHGYEYPTSGRRYQGTGICITFGYSQFEEDFOLL 180  
 QY 181 VLOPNYSCFIPSTFHWKLIYMPRMCANSSSLKIPVRYLAELHSAOGLNFVHFAKSSFYT 240  
 Db 181 VLOPNYSCFIPSTFHWKLIYMPRMCANSSSLKIPVRYLAELHSAOGLNFVHFAKSSFYT 240  
 QY 241 DDIFTGWIAQKLTHLLAQTWKKQELPNCSLPYHVNIKSGIVTSKSYFSSRDHSHK 300  
 Db 241 DDIFTGWIAQKLTHLLAQTWKKQELPNCSLPYHVNIKSGIVTSKSYFSSRDHSHK 300  
 QY 301 WCVSIKGSANRWTCIGDLNRSLSHQALRGGGFICTKNHYIYQAFHKLYLRYGFCCK 354  
 Db 301 WCVSIKGSANRWTCIGDLNRSLSHQALRGGGFICTKNHYIYQAFHKLYLRYGFCCK 354

RESULT 3  
 AAE13015  
 ID AAE13015 standard; protein; 357 AA.  
 AC AAE13015;  
 XX  
 XX 28-JAN-2002 (first entry)  
 DT  
 DE Human deoxyribonuclease (DNase) II beta protein.  
 XX  
 XX Human; deoxyribonuclease; DNase II beta protein; mucous plug; lung;  
 KW cystic fibrosis; pulmonary; chromosome 1p22.  
 XX  
 XX Homo sapiens.  
 XX  
 XX WO200175082-A1.  
 XX  
 XX 11-OCT-2001.  
 XX  
 XX 02-APR-2001; 2001WO-US010635.  
 XX  
 XX 03-APR-2000; 2000US-00541840.  
 XX  
 XX 19-MAY-2000; 2000US-00574942.  
 XX  
 XX (DART-) DARTMOUTH COLLEGE.  
 XX  
 XX Eastman AR, Krieser RJ;  
 XX  
 XX WPI; 2001-662972/76.  
 XX  
 XX N-PSDB; AAD21289.  
 XX  
 XX New cDNA encoding a deoxyribonuclease II beta enzyme useful for degrading  
 PT DNA present in the mucous plugs in the lungs of cystic fibrosis patients.  
 XX  
 XX Claim 5; Page 16-17; 21pp; English.  
 XX  
 XX The invention relates to deoxyribonuclease (DNase) II beta proteins and  
 CC their corresponding cDNAs. The DNase II beta may be useful to digest DNA  
 CC in the mucous plugs in lungs of cystic fibrosis patients and so reduce  
 CC their viscosity. The present sequence is human DNase II beta protein. The  
 CC human DNase II beta gene is located at chromosome 1p22  
 XX  
 XX Sequence 357 AA;  
 SQ  
 Query Match 67.1%; Score 1293.5; DB 4; Length 357;  
 Best Local Similarity 66.4%; Pred. No. 4.8e-125;  
 Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

QY 1 MTAKPLRTVLSLLFFALSGVLGTPETSCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYL 60

Db 1 MWARLRTSFGALLFLGLFGLVGAATISCRNEEGKAVDMFTFYKLPKQNKESGETGLEVL 60  
Qy 61 YLDSTQWTKNSLYLINSRSGALGRTLOHLYDTHNS-TNDTAYLIYNDGVPQSVNSROY 119  
Db 61 YLDSTTRSRKSEQLMNDTKSVLGRTOLOLYEAYASKSNNTAYLIYNDGVPKPVNSRYK 120  
Qy 120 GHAKGLLVNRTQGFGLHSHVDPKPPV--HGVEYPTSGRRYGTGICITFGYSQPEEIDF 177  
Db 121 GHTKGLLVNRTQGFGLHSHVDPKPPV--HGVEYPTSGRRYGTGICITFGYSQPEEIDF 180  
Qy 178 QLLVLPQNIYSCFIPSTFHWKLIYMPRCANSSSLKIPVRYLAELHSAQGLNFVHFAKSS 237  
Db 181 QLLVLPQNIYSCFIPSTFHWKLIYMPRCANSSSLKIPVRYLAELHSAQGLNFVHFAKSD 240  
Qy 238 FYTDDIFTGWAQKLKTHLLAQTWQKKQELPSNCSLPVHVYVNIKSIQVTSKSYFSSROD 297  
Db 241 SFLLDDIFAAMWAQRLKTHLLTETWQKQELPSNCSLPVHVYVNIKSIQVTSKSYFSSYOD 300  
Qy 298 HSKWCYSIKGSANRWTCIGDLNRSIHOALRGGFICTKNHVIYQAFHKLILRYGCK 354  
Db 301 HAKWCISQKGTNRWTCIGDLNRSIHOALRGGFICTKNHVIYQAFHKLILRYGCK 357

RESULT 4  
ID ABG61821 standard; protein; 357 AA.  
AC ABG61821;  
XX

DT 15-AUG-2002 (first entry)  
XX

DE Prostate cancer-associated protein #22.  
XX

DE Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.  
XX

OS Mammalia.  
XX

PN WO200230268-A2.  
XX

PD 18-APR-2002.  
XX

PF 12-OCT-2001; 2001WO-US032045.  
XX

PR 13-OCT-2000; 2000US-00687576.  
XX

PR 08-DEC-2000; 2000US-00733298.  
XX

PR 08-DEC-2000; 2000US-00733742.  
XX

PR 24-JAN-2001; 2001US-0263957P.  
XX

PR 16-MAR-2001; 2001US-0276791P.  
XX

PR 06-APR-2001; 2001US-0276888P.  
XX

PR 24-APR-2001; 2001US-0281922P.  
XX

PR 30-APR-2001; 2001US-0286214P.  
XX

PR 04-MAY-2001; 2001US-0288589P.  
XX

PA (EOSB-) EOS BIOTECHNOLOGY INC.  
XX

PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;  
XX

XX WPI; 2002-471335/50.  
XX

DR N-PSDB; ABK92136.  
XX

PT Detecting a prostate cancer-associated transcript in a cell in a patient,  
XX

PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,  
XX

PT by determining if prostate cancer-associated genes are expressed in a  
XX

PT prostate tissue.  
XX

PS Claim 27; Page 317; 436pp; English.  
XX

XX The present invention relates to methods of detecting a prostate cancer-  
XX

CC hybridise to a sequence that is at least 80% identical to them. The  
CC prostate cancer-associated polynucleotide sequences are differentially  
CC expressed in prostate tumour tissue or in prostate cancer and are derived  
CC from the tissues of various organisms such as humans or other mammals  
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for  
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-  
CC associated genes are useful for diagnosing or treating prostate cancer,  
CC as well as for identifying modulators of prostate cancer or agents that  
CC inhibit prostate cancer. The nucleic acid sequences are particularly  
CC useful in gene therapy, as a vaccine or in antisense applications.  
CC ABG61800-ABG61944 represent prostate cancer-associated proteins  
XX  
SQ Sequence 357 AA;

Query Match 67.1%; Score 1293.5; DB 5; Length 357;  
Best Local Similarity 66.4%; Pred. No. 4.8e-125;  
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

Qy 1 MTAKPLRTVLSLLFPALSGVLGTPRISCRNEEGKAVDMFTFYKLPKQNKESGETGLEVL 60  
Db 1 MWARLRTSFGALLFLGLFGLVGAATISCRNEEGKAVDMFTFYKLPKQNKESGETGLEVL 60  
Qy 61 YLDSTQWTKNSLYLINSRSGALGRTLOHLYDTHNS-TNDTAYLIYNDGVPQSVNSROY 119  
Db 61 YLDSTTRSRKSEQLMNDTKSVLGRTOLOLYEAYASKSNNTAYLIYNDGVPKPVNSRYK 120  
Qy 120 GHAKGLLVNRTQGFGLHSHVDPKPPV--HGVEYPTSGRRYGTGICITFGYSQPEEIDF 177  
Db 121 GHTKGLLVNRTQGFGLHSHVDPKPPV--HGVEYPTSGRRYGTGICITFGYSQPEEIDF 180  
Qy 178 QLLVLPQNIYSCFIPSTFHWKLIYMPRCANSSSLKIPVRYLAELHSAQGLNFVHFAKSS 237  
Db 181 QLLVLPQNIYSCFIPSTFHWKLIYMPRCANSSSLKIPVRYLAELHSAQGLNFVHFAKSD 240  
Qy 238 FYTDDIFTGWAQKLKTHLLAQTWQKKQELPSNCSLPVHVYVNIKSIQVTSKSYFSSROD 297  
Db 241 SFLLDDIFAAMWAQRLKTHLLTETWQKQELPSNCSLPVHVYVNIKSIQVTSKSYFSSYOD 300  
Qy 298 HSKWCYSIKGSANRWTCIGDLNRSIHOALRGGFICTKNHVIYQAFHKLILRYGCK 354  
Db 301 HAKWCISQKGTNRWTCIGDLNRSIHOALRGGFICTKNHVIYQAFHKLILRYGCK 357

RESULT 5  
AAB72417  
ID AAB72417 standard; protein; 361 AA.  
XX  
AC AAB72417;  
XX

DT 03-MAY-2001 (first entry)  
XX

DE Human DNase.  
XX

XX Human; DNase; DLAD; deoxyribonuclease; endonuclease; cystic fibrosis;  
XX infectious disease.  
XX

OS Homo sapiens.  
XX

XX WO200112793-A1.  
XX

PD 22-FEB-2001.  
XX

XX 01-MAY-2000; 2000WO-JP002893.  
XX

XX 17-AUG-1999; 99JP-00230870.  
XX

XX (TANU/) TANUMA S.  
XX

XX Tanuma S, Shiokawa D;  
XX

XX WPI; 2001-218348/22.  
XX

XX N-PSDB; AAF60716.  
XX

PT Acidic deoxyribonuclease capable of divalent cation-independent cleavage  
 PT of DNA under acidic even neutral pH and not inhibited by G-actin, useful  
 PT in remedies for cystic fibrosis and for treatment of infectious diseases.

XX Claim 8; Page 52-53; 61pp; Japanese.

XX The present sequence is a human deoxyribonuclease (DLAD), which is an  
 CC endonuclease. DLAD is capable of divalent cation-independent cleavage of  
 CC DNA under acidic conditions. This protein can be used as a substitute for  
 CC DNase I in treating cystic fibrosis, and is useful in the prevention and  
 CC treatment of infectious diseases

XX Sequence 361 AA;

Query Match 66.8%; Score 1286.5; DB 4; Length 361;  
 Best Local Similarity 66.1%; Pred. No. 2.6e-124;  
 Matches 236; Conservative 45; Mismatches 73; Indels 3; Gaps 2;  
 QY 1 MTAQRLTVLSLFFALSGVLGTPPEISCRNEYGEAVDWFIFKLPKRTSKASEAGLQYL 60  
 Db 5 MWAQLRTSFALLFLGFGVLGATISCRNEEGKAVDWFIFKLPKQNKESGETGLYL 64  
 QY 61 YLDSTRTQWNSLYLINSTSALGRTLOHLYDPHNS-TNDTAYLIYNDGVPGSVNSROY 119  
 Db 65 YLDSTTRSWEKSQLMMDTKSVLGRTLQQLYEAYASKSNNTAYLIYNDGVPKPNYSRKY 124  
 QY 120 CHAKGLLVNRTQFWLIHVPKFPV--HGVEYPTSGRYGOTGICITFGYQFEIDF 177  
 Db 125 GHTKGLLVNRTQFWLIHVPKFPV--HGVEYPTSGRYGOTGICITFGYQFEIDF 184  
 QY 178 QLVLPQNIYSCIPSTFHWKLIYMPKNCANSSSLKIPVRYLAELHSAQGLNFWHFAKSS 237  
 Db 185 QLVLPQNIYSCIPSTFHWKLIYMPKNCANSSSLKIPVRYLAELHSAQGLNFWHFAKSS 244  
 QY 238 FYTDDIFTGIAQKLTLLAQTWQKKQBLPNCNLSLPHVYNIKISGVTSKYFSSRD 297  
 Db 245 SFLDGIFAAWMAQKLTLLAQTWQKKQBLPNCNLSLPHVYNIKISGVTSKYFSSRD 304  
 QY 298 HSKWCYSIKSANRTWICIGLNLSLHQAALRGGFICTKNHYIYQAFHKLRYGFC 354  
 Db 305 HAKWCISQKTKNRTWICIGLNLSLHQAALRGGFICTKNHYIYQAFHKLRYGFC 361

RESULT 6  
 AAW31495  
 ID AAW31495 standard; protein; 360 AA.  
 XX AAW31495;  
 AC  
 DT 11-MAY-1998 (first entry)  
 XX Human DNase II protein.  
 DE  
 XX DNase II; human; deoxyribonuclease; pulmonary disease; diagnostic;  
 KW cellular DNA debris.  
 KW Homo sapiens.

XX Key Location/Qualifiers  
 FH Peptide 1..16  
 FT /label= signal\_peptide  
 FT Protein 17..360  
 FT /note= "DNase II"  
 FT Modified-site 86..88  
 FT /note= "Asn is N-glycosylated"  
 FT Modified-site 212..214  
 FT /note= "Asn is N-glycosylated"  
 FT Modified-site 266..268  
 FT /note= "Asn is N-glycosylated"  
 FT Modified-site 290..292  
 FT /note= "Asn is N-glycosylated"

WO9740134-A2.

XX 30-OCT-1997.  
 PD 23-APR-1997; 97WO-US006664.  
 XX 25-APR-1996; 96US-00639294.  
 PR (GETH ) GENENTECH INC.  
 PA Baker KP, Baron WF;  
 XX WPI; 1997-535820/49.  
 DR N-PSDB; AAV02903.  
 XX Human DNase II and related nucleic acids - useful in protein production  
 PT e.g. for therapeutic use to treat systemic lupus erythematosus and  
 PT pulmonary diseases e.g. cystic fibrosis.  
 XX Claim 1; Fig 1; 30pp; English.  
 PS This sequence represents a novel human deoxyribonuclease, DNase II. This  
 CC protein is useful to reduce the viscoelasticity of DNA-containing  
 CC material, e.g. mucus, and used to treat patients with pulmonary diseases  
 CC or disorders e.g. cystic fibrosis, or with systemic lupus erythematosus.  
 CC It may also improve antibiotic efficacy in the treatment of abscesses,  
 CC infected lesions etc., provide treatment in non-infected conditions in  
 CC which there is an accumulation of cellular DNA debris, e.g.  
 CC pyelonephritis and be used to degrade DNA in biological samples or in  
 CC diagnostic assays. The nucleic acids and vectors can be used for in vivo  
 CC or ex vivo gene therapy, and antisense oligonucleotides can be produced  
 CC from the nucleic acids which can bind to and prevent expression of  
 CC nucleic acid within cells. The nucleic acids (or a portion) can also be  
 CC used for hybridisation assays for nucleic acids encoding human DNase II  
 CC in a sample, or to identify and isolate nucleic acids sharing substantial  
 CC sequence identity (e.g. encoding naturally-occurring allelic variants of  
 CC human DNase II). The antibodies can be used to detect and measure human  
 CC DNase II in tissues or clinical samples, and in the purification of human  
 CC DNase II. The nucleic acids enable production of human DNase II by  
 CC recombinant DNA methods in quantities sufficient for clinical use, not  
 CC previously possible  
 XX Sequence 360 AA;

Query Match 31.0%; Score 596.5; DB 2; Length 360;  
 Best Local Similarity 38.0%; Pred. No. 9.8e-53;  
 Matches 132; Conservative 56; Mismatches 144; Indels 15; Gaps 7;  
 QY 17 LSGVLGTP--EISCRNEYGEAVDWFIFKLPK-RTSKASEAGLQYLIDSTRTQWNSKL 73  
 Db 6 LAALLCVPAGALTGYDGSQGVDFVYVYKLPALRGSGEAAQRLQYKYLDESSGGWRDGR 65  
 QY 74 YLINSTSALGRTLOHLYDTHNSTNDTAYLIYND--GVPGSVNSROYGHAKGLLVNRT 131  
 Db 66 ALINSPGAVGRSLQPLY--RSNTSLQALFLYNDQPPQPSKAQDSSNRGKTKGVLLDHD 123  
 QY 132 QGFWLHSPKFPFP---VHGVEYPTSGRYGOTGICITFGYQFEIDFOLLQPNYS 189  
 Db 124 GGFWLHSPVNFPPPPASSAAYSWPHSACTYQGLLQVSPFPFAQFSKMGKQLTYTPWVYN 183  
 QY 189 CFIPSTFWKLIYMPKNCANSSSLKIPVRYLAELHSAQGLNFWHFAKSSFYTDIFGWI 248  
 Db 184 YQLEGIPFAQFPDLENVNVKGGHVSQEPWNSITLTSQAGAVFQSFQKFGDDLVSGLW 243  
 QY 249 AQKLTLLAQTWQKKQBLPNCNLSLPHVYNIKISGVTSK--YFSSRODHSKWCYSIK 306  
 Db 244 AALGTNLQVFWHKTGVLIPSCNLSLQVNLVNVQIAFPAGPAGPSFNSHDKWCYSPK 303  
 QY 307 GSANRTWICIGLNLSLHQAALRGGFICTKNHYIYQAFHKLRYGFC 353  
 Db 304 GP---WTCVGMNRRNGEQRGGGTLCALPALWKAQPLVKNYQPC 347

RESULT 7



AAW37920  
ID AAW37920 standard; protein; 360 AA.  
AC AAW37920;  
XX  
XX  
DT 11-SEP-1998 (first entry)  
DE Homo sapiens DNase II.  
XX  
XX DNase II; deoxyribonuclease II; cystic fibrosis; DNA digestion;  
KW controlled cell death; apoptosis; metaplasia; cell turnover;  
KW neurodegeneration; AIDS; acquired immune deficiency syndrome; cancer;  
KW autoimmune disease; diagnosis; anticancer agents;  
KW enhanced chromosomal rearrangement; chromosome instability.  
XX  
XX Homo sapiens.  
XX  
XX WO9816659-A1.  
FN 23-APR-1998.  
XX  
XX 09-OCT-1997; 97WO-US018262.  
XX  
XX 15-OCT-1996; 96US-0028539P.  
PR (DART-) DARTMOUTH COLLEGE.  
PA Eastman A, Krieser R;  
PI WPI; 1998-251301/22.  
DR N-PSDB; AAV29137.  
XX  
XX Human and bovine deoxyribonuclease II enzyme and encoding cDNA - useful,  
PT e.g. to induce apoptosis in tumour cells, identify agents modulating  
PT apoptosis and digest DNA in sputum of cystic fibrosis patients.  
XX  
XX Claim 5; Page 18; 29pp; English.  
XX  
XX The sequence is that of deoxyribonuclease II (DNase II). The enzyme is  
CC useful to digest DNA, e.g. in the lung sputum of cystic fibrosis patients  
CC to reduce sputum viscosity. It (or fragments) are also useful to produce  
CC antibodies, e.g. to study DNase II expression in cells. DNase II has  
CC recently been linked with DNA fragmentation in the early stages of  
CC controlled cell death (apoptosis), a process critical to homeostasis  
CC during, e.g. metamorphosis or cell turnover. Too much cell death can lead  
CC to neuro-degeneration and acquired immune deficiency syndrome (AIDS),  
CC whilst too little can lead to cancer or autoimmune diseases. The  
CC antibodies raised against it can therefore be used to diagnose apoptotic  
CC stages in selected cells, by contacting cells with the antibody,  
CC detecting binding of the antibody with DNase II and determining DNase II  
CC levels. Such diagnosis is useful to evaluate the efficacy of therapeutic  
CC agents, e.g. anticancer agents to promote apoptosis in cells. The  
CC antibodies can also be used to identify cells susceptible to premature  
CC death. The cDNA encoding it is useful to identify agents modulating  
CC apoptosis in cells, by treating cells with an agent, transfecting cells  
CC with cDNA and monitoring apoptosis compared with untreated cells.  
CC Inhibitors identified may be useful in preventing diseases relating to  
CC enhanced chromosomal re-arrangement. Vectors comprising the cDNA can be  
CC used to induce apoptosis in selected cells, e.g. tumour cells or cells  
CC involved in autoimmune disorders. Antisense oligonucleotides can be  
CC administered to cells to inhibit DNase II expression to reduce chromosome  
CC instability associated with cancer  
XX  
XX Sequence 360 AA;

Query Match 31.0%; Score 596.5; DB 2; Length 360;  
Best Local Similarity 38.0%; Pred. No. 9.8e-53;  
Matches 132; Conservative 56; Mismatches 144; Indels 15; Gaps 7;  
XX 17 LSGVLGTP--EISCNEYGEAVDWIFKLPK-RYSKASEAGLOLYLDSTRQWNSL 73  
DB 6 LAALLCVFAGALTCTCGDGGQPDWFFVVKLPALRGSGEAAQRLQYKYLDSSGGWRDGR 65

QY 74 YLINSTRSALGRTLOHLYDTHNSTNDTAVLIYND--GVPSVNYSPQYGHAKGLVWNR 131  
DB 66 ALINSPGAVGRSLQPLY--RNTSOLAFLLYNDQPPQSKAODSSMRGHTKGLVLLDHD 123  
QY 132 QGFVLIHSPKFP--VHGVEYPTSGRRYGTGICITFGYSQFEIDFOLLVLPQNIYS 188  
DB 124 GGFVLIHSPVNPFPASSAAYSWPHSACTYGTLLCVSPFPAQFSKMGKQLTYTPWVYN 193  
QY 189 CRIPSTFWKLLYMPRMCMANSSSLKIPVRYLAELHSAOGLNFVHFAKSSFYDDIFTGWI 248  
DB 184 YQLEGIFAQEPDLENVVKGHVSOEPMNSITLTSQAGAVFQSPAKFSKFGDLYSGWL 243  
QY 249 AQKLTALLAQWQKKQELPNSCSIPVHYNIKISGVTSKS--YESSRQDSKVCVSTK 306  
DB 244 AALGTNLQVFWHTVGLPNSCSDIWOVLNVNQAAPPAGPSFNSFTHEDSKVCVSPK 303  
QY 307 GSANRWTCIGDLNRLHQAALRGSGFICTKNHYIYQAFHKLRLRYGFC 353  
DB 304 GP--WTCVGMNRNQGEQRGGGTLCAQLPALWKAQFPLVKNYQPC 347  
XX  
XX ABB63496 standard; protein; 366 AA.  
AC ABB63496;  
XX  
XX 26-MAR-2002 (first entry)  
XX  
XX Drosophila melanogaster polypeptide SEQ ID NO 17280.  
DE Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
XX Drosophila melanogaster.  
XX  
XX WO200171042-A2.  
XX  
XX 27-SEP-2001.  
XX  
XX 23-MAR-2001; 2001WO-US009231.  
XX  
XX 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX (PEKE ) PE CORP NY.  
XX  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX  
XX WPI; 2001-656860/75.  
DR N-PSDB; ABL07599.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions.  
XX  
XX Disclosure; SEQ ID NO 17280; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
CC sequences (AB101840-AB16175) and the encoded proteins (ABBS7737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_sequences  
XX  
XX Sequence 366 AA;

Query Match 23.0%; Score 444; DB 4; Length 366;  
Best Local Similarity 31.0%; Pred. No. 6.6e-37;



```
XX PR      11-DEC-1997;   97US-006920SP.  
XX PA      (REGC ) UNIV CALIFORNIA.  
XX PI      Aguilera RJ, Lyon CJ;  
XX XX  
DR WPI; 2003-065229/06.  
DR N-PSDB; ABS58030.  
XX XX  
PT New nucleic acid and its encoded murine endonuclease-SR protein, useful  
FT for treating tumors or cancers, e.g. leukemias or lymphomas, or for  
PT designing and isolating peptidomimetics or inhibitors of Endo-SR for  
ET treating these diseases.  
XX XX  
PS Example 5; Col 82; 66pp; English.  
XX XX  
CC The invention discloses an isolated nucleic acid molecule, which encodes  
CC a murine endonuclease implicated in switch recombination (Endo-SR)  
CC protein. The enzyme is useful at cleaving DNA at specific G-rich regions  
CC which are implicated in modulating DNA rearrangements. Also disclosed are  
CC methods for repairing DNA and modulating genetic recombination in a cell.  
CC One example of genetic rearrangement is in mature B lymphocytes which can  
CC alter their Ig constant region for another, significantly enhancing the  
CC versatility of the immune system by allowing B cells to alter their  
CC function without altering their ligand specificity. The endo-SR nuclear  
CC acid and protein are useful (using gene therapy) for treating tumours or  
CC cancers, e.g. leukaemias or lymphomas, as well as the modulation of  
CC apoptosis and programmed cell death events. The endo-SR nuclear acid and  
CC protein are also useful for designing and isolating peptidomimetics and  
CC inhibitors of Endo-SR, which may be employed for treating the diseases.  
CC The sequence presented is the mouse endo-SR protein  
XX XX  
SQ Sequence 192 AA;
```

```
Query Match           12.3%; Score 236.5; DB 6; Length 192;  
Best Local Similarity 27.1%; Pred. No. 8.5e-16;  
Matches    67; Conservative    27; Mismatches    96; Indels     57; Gaps       4;
```

```
QY          6 LRTVLSLLFFALSGVLGTPETSICRNVEGEAVDWFIFYKLPKRTSKAEEAGLQLYLIDST 65  
Db         : |::|::||::|::|::|::|::|::|::|::|::|::|::|::|:  
            MATURLSLLAALLWV-PAAELSCYGDSGPVDWEFVVKLFAHSGSRDPKGLTCKYMDDQN 59
```

```
QY        66 RQTWNKSLEYINSTRSALGRTLQHLYPTHNSTNDTAYLIYNDGVFP--GSVNYSRYQGHA 123  
Db         : ||::|::|::|::|::|::|::|::|::|::|::|::|::|:  
            SDGWQQDGVGYNISPEGAVERSLQPPLY--RKNSSQLAFLLYNDQPKSSSTDSTGHGHTK 117
```

```
QY        124 GLLVNVRTQGFLLHSVPKPFPVHGIVEYPFSRGRRYTGTCTITFGYQSQFEIIDLVLVLQ 183  
Db         118 GK-----  
            PNITYSCFIPTSFHWKLIYMPRMCAANSSSLKIPIVRYLELAELSACGLNFVFHAKSSFYTD 243
```

```
Db        126 PLYDYHKLEGGFAKKPLDLDTVIKNQHVLFHEPWNSSVLITSQAATQSFSAKFQKFGDDL 185
```

```
QY        244 FTGNIIAQ 250  
Db         ::||::|:  
            YSGWLAE 192
```

```
RESULT 11  
RAY29333 standard; protein; 475 AA.
```

```
XX AC AAY29333,  
XX DT 29-SEP-1999 (first entry)
```

```
XX DE Human secreted protein clone pj323_2 protein sequence.
```

```
XX KW Human; secreted protein; nutrition; cytokine; cell proliferation;  
KW differentiation; immune stimulating; vaccine; suppression; gene therapy;  
KW haematopoiesis regulation; tissue growth; activin; inhibin csadherin;
```

KW	chemotactic; chemokinetic; haemostatic; thrombolytic; anti-inflammatory;
KW	tumour invasion suppressor; tumour inhibition.
XX	
XX	Homo sapiens.
OS	XX
XX	XX
PN	W09937674-A1.
XX	XX
PD	29-JUL-1999.
XX	XX
PF	21-JAN-1999; 99WO-US001404.
XX	XX
PR	22-JAN-1998; 98US-0072134P.
PR	20-JAN-1999; 99US-00235609.
XX	XX
PA	(GEMY ) GENETICS INST INC.
XX	XX
PI	Jacobs K, Mccoy JM, Lavallie ER, Collins-Racie LA, Merberg D;
PI	Tracy M, Agostino MJ, Steininger RJ, Wong GG, Clark HF, Fechtel K;
XX	XX
DR	WPI: 1999-458682/38.
DR	N-PSDB; AAX30448.
XX	XX
PT	New polynucleotides encoding secreted human proteins derived from, e.g.
PT	fetal brain potentially used as immunostimulators.

Claim 34; Page 126-128; 139pp; English.

The present sequence represents a human secreted protein. Human secreted protein polynucleotides and proteins are predicted to have biological activities which would make them suitable for treating, preventing or ameliorating medical conditions in humans and animals, although no supporting data is given. Suggested activities include nutritional activity, cytokine and cell proliferation/differentiation activity, immune stimulating (e.g. as vaccines) or suppressing activity, haematopoietic regulating activity, tissue growth activity, actinin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, receptor/ligand activity, anti-inflammatory activity, cadherin/tumour invasion suppressor activity, and tumour inhibition activity. The polynucleotides are also stated to be useful for gene therapy

Sequence 475 AA;  
SQ

Query Match	5.3%; Score 101.5; DB 2; Length 475;
Best Local Similarity	21.4%; Pred. No. 0.33;
Matches	100; Conservative 48; Mismatches 147; Indels 173; Gaps 23;

  

QY	6	LRTVLSLLPFA	---	LSGVLTGTEISCRNBYGRAWDF	-----	IFYKLPKRTSKASE	53
		:		:		:	
Db	40	LASVIFLLFAPFIVYVFIMACDQYSALT	-	GPVVDIVTGHARLSDIWAKTPPTTRKAAQ	98		
		:		:		:	
QY	54	EAGI	---	OZVLDSRQTWNKSL	-----	YLNSTFESALGRLOH	89
Db	99	LYTLWTFTQVLLYTSIPDFCKFLPGVVGIGQEGAVTPAGVANKYQINGLOAMLLTHLLW	158				
QY	90	LYDTHNSTNDYALIYNDGVPGSVNYSRQYGHAKGLLVNMNRTQGFMLHISVXPFPVPHGY	149				
		:		:		:	
Db	159	FANAHLLSWFSPIIIFDNWIP	-----	LLWCAN	--	ILGYAVSTFAMVKGY	200
QY	150	EYPTSGRRYQGTGICITFGYSQPERIDFOLLVLQPNYSCF	-----	IPSTFHWKLIY	201		
Db	201	FFPETSARDCKFTG	--	NFFNYVMGIEF	---	NPRIGKWDFKLFNGRPGIVAWTLIN	252
QY	202	MPRMCANSLSLIPVYLAELHS	-	AQGLNFVHFVFAKSSPYTDDIFGTGWAQKLKTHLLAQ	259		
		:		:		:	
Db	253	L	-----	SFAAK	-----	QRELHSHVTNAMVLNVLO	---NE 288
QY	260	TWOKKQOEL	-----	PSNCSLP	-----	Y	276
Db	289	TWYLKIDICHDFHGLGWGDCWLPYLYTQLGLYLVTHVPOLSTPHAVGVLLGLNGVY	348				
QY	277	HVYNI	-----	KSIGVTSKSYFSS	--	RQDHSKWCVS	IKGSANRWTC 314
		:		:		:	

Db	349	YIFRVANHQDLFRRTDGRCLWGRPKVIECSYTSADGQRHSHKLLVSGFWGVARHFN	408
QY	315	IGDLNLSLHQALR-GGGFICTKNHYIQAF-----HKLYLRG	351
Db	409	VGDLMSLAYCLACGGGHLFPYIIYMAILLTHRCLRDEHRCASKYG	456
RESULT 12			
AAW93573	ID AAW93573 standard; protein; 475 AA.		
XX	AC	AAW93573;	
XX	DT	17-JUN-1999 (first entry)	
XX	DE	Human Delta7-sterol reductase protein.	
XX	KW	Human; Delta7-sterol reductase; hereditary; Smith-Lemli-Opitz syndrome;	
XX	KW	diagnosis; screening; double bond removal; 7-dehydrocholesterol;	
XX	KW	organic polymeric ring; cholesterol.	
XX	OS	Homo sapiens.	
XX	PN	DE19739940-A1.	
XX	PD	18-MAR-1999.	
XX	PF	11-SEP-1997; 97DE-01039940.	
XX	PR	11-SEP-1997; 97DE-01039940.	
XX	PA	(GLOS/) GLOSSMANN H.	
XX	PI	Glosemann H, Moebius F, Fitzky B;	
XX	XX	WPI; 1999-191430/17.	
XX	DR	N-PSDB; AAX23387.	
XX	PT	Human Delta7-sterol reductase polypeptide - useful for diagnosis or	
XX	PT	treatment of genetic defects e.g. hereditary Smith-Lemli-Opitz syndrome.	
XX	PS	Disclosure; Page 29-30; 62pp; German.	
XX	CC	This invention describes a human Delta7-sterol reductase. The encoding	
XX	CC	DNA can be used to diagnose or correct human Delta7-sterol reductase gene	
XX	CC	defects e.g. hereditary Smith-Lemli-Opitz syndrome, and to produce the	
XX	CC	recombinant Delta7-sterol reductase polypeptide, which can be used to	
XX	CC	replace a defective Delta7-sterol reductase enzyme in humans or other	
XX	CC	animals. It is also useful to screen for Delta7-sterol reductase	
XX	CC	inhibitors or to introduce and remove double bonds in synthetic and	
XX	CC	naturally occurring organic polymeric ring systems (Delta 7-sterol	
XX	CC	reductase catalyses the conversion of 7-dehydrocholesterol to	
XX	CC	cholesterol)	
XX	XX	Sequence 475 AA;	
Query Match			
Best Local Similarity 5.3%; Score 101.5; DB 2; Length 475;			
Matches 100; Conservative 48; Mismatches 147; Indels 173; Gaps 23;			
QY	6	LRTVLSLFFA---LSGLGTPTISCRNEYGEADVWF-----IFYKLPRKTSKASE	53
Db	40	LASVIFLLFAPFIVYVIMACDQYSCALT-GPVVDIVTGHARLSDIWAKTPTIRKAAQ	98
QY	54	EAGL-----QYILSTQTNKSL-----YLNSTRSALGRILQH	89
Db	99	LYTLWVTFQVLLYTSLPDFCHKFLPGYGGIPEGAVTPAGVVKYQINGLQALLTHLLW	158
QY	90	LYDTHSTNDTAXLYLNDGVPSVNSQYGHAKGLLVNWNRTQGFLLIHSVPKPPVHG	149
Db	159	FANAHLLSWFPTIIFDWNIP-----LLWCAN---ILGYAVSTAMVKGY	200
QY	150	EYPTSGRRYGTGICITFGYSQFEIDFQLLVLPQNIYSCF-----IPSTFWKLIY	201



RESULT 15  
 ABB55768  
 ID ABB55768 standard; protein; 475 AA.  
 AC ABB55768;  
 XX  
 DT 14-FEB-2002 (first entry)  
 XX  
 DE Human polypeptide SEQ ID NO 142.  
 XX  
 KW Human; clone b306-7; ATCC number 98599; gene therapy;  
 KW immune disorder; bacterial infection; fungal infection; cancer; tumour;  
 KW autoimmune disorder; systemic lupus erythematosus; wound; ulcer; inhibin;  
 KW osteoporosis; osteoarthritis; nervous system disorder; neuropathy;  
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; activin;  
 KW haemophilia; cardiac infarction; stroke; sepsis; arthritis; vulnery;  
 KW ischaemia-reperfusion injury; inflammatory bowel disease; chemotactic;  
 KW Crohn's disease; cytostatic; anti-inflammatory; immunomodulator;  
 KW neuroprotective; haemostatic; thrombolytic; anti-inflammatory.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2001039335-A1.  
 XX  
 PD 08-NOV-2001.  
 XX  
 PF 04-DEC-2000; 2000US-00729674.  
 XX  
 PR 26-NOV-1997; 97US-0126425P.  
 PR 04-DEC-1997; 97US-0067454P.  
 PR 20-DEC-1997; 97US-0068379P.  
 PR 02-JAN-1998; 98US-0070346P.  
 PR 07-JAN-1998; 98US-0070643P.  
 PR 08-JAN-1998; 98US-0070755P.  
 PR 13-JAN-1998; 98US-0071304P.  
 PR 22-JAN-1998; 98US-0072134P.  
 PR 30-JAN-1998; 98US-0073095P.  
 PR 18-FEB-1998; 98US-0075038P.  
 PR 23-NOV-1998; 98US-00197886.  
 PR 30-MAR-2000; 2000US-00539330.  
 XX  
 PA (JACO/) JACOBS K.  
 PA (MCCO/) MCCOY J M.  
 PA (LAVA/) LAVALLIE R R.  
 PA (COLL/) COLLINS-RACIE L A.  
 PA (EVAN/) EVANS C.  
 PA (MERB/) MERBERG D.  
 PA (TREA/) TREACY M.  
 PA (AGOS/) AGOSTINO M J.  
 PA (STEI/) STEININGER R J.  
 PA (SPAU/) SPAULDING V.  
 PA (WONG/) WONG G G.  
 PA (CLAR/) CLARK H.  
 PA (FECH/) FECHTEL K.  
 XX  
 PI Jacobs K, Mccoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
 PI Merberg D, Treacy M, Agostino MJ, Steininger RJ, Spaulding V;  
 PI Wong GG, Clark H, Fechtel K;  
 XX  
 DR WPI; 2002-040725/05.  
 DR N-PSDB; ABA90946.  
 XX  
 PT New secreted proteins and encoding polynucleotides, useful in gene  
 PT therapies, particularly for preventing or treating autoimmune disorders,  
 PT cancer, graft-versus-host disease, wound, osteoporosis, stroke or  
 PT inflammations.  
 XX  
 PS Disclosure; Page 281-282; 349pp; English.  
 XX  
 CC The invention relates to isolated polynucleotides (ABA90876-ABA90968 and  
 CC ABA90980) and encoded proteins (ABB55698-ABB55800), especially  
 CC polynucleotides SEQ ID NO 1 (ABA90876) and SEQ ID NO 19 (ABA90885) and

CC proteins SEQ ID NO 2 (ABB55698) and SEQ ID NO 20 (ABB55707) contained in  
 CC clones b306-7 and yb8-1 respectively and the clones b306-7 and yb8-1  
 CC are deposited with the American Type Culture Collection (ATCC) with  
 CC accession number 98599. The polynucleotides and encoded polypeptides have  
 CC cytotactic, anti-inflammatory, immunomodulator, vulnerary,  
 CC neuroprotective, activin, inhibin, chemotactic, haemostatic, thrombolytic  
 CC and anti-inflammatory activity and acting as cytokine modulators,  
 CC haematopoiesis regulators, tissue growth modulators and/or cadherin  
 CC suppressors. The polypeptides and polynucleotides are useful in gene  
 CC therapies, particularly for preventing, treating or ameliorating any of  
 CC the following diseases: immune deficiency and disorders; e.g. bacterial  
 CC or fungal infections, autoimmune disorders, cancer, systemic lupus  
 CC erythematosus or graft-versus-host disease; myeloid or lymphoid cell  
 CC deficiencies; wound, burns, incisions and ulcers, osteoporosis or  
 CC osteoarthritis; central and peripheral nervous system diseases and  
 CC neuropathies, e.g. Alzheimer's, Parkinson's disease, Huntington's  
 CC disease, amyotrophic lateral sclerosis or Shy-Drager syndrome;  
 CC haemophilia, cardiac infarction or stroke; inflammations, shock, sepsis  
 CC or systemic inflammatory response syndrome, ischaemia-reperfusion injury,  
 CC endotoxin lethality, arthritis, inflammatory bowel disease or Crohn's  
 CC disease; or tumours or cancers, pemphigus vulgaris or pemphigus foliaceus  
 XX  
 SQ Sequence 475 AA;  
 Query Match 5.3%; Score 101.5; DB 5; Length 475;  
 Best Local Similarity 21.4%; Pred. No. 0.33;  
 Matches 100; Conservative 48; Mismatches 147; Indels 173; Gaps 23;  
 QY 6 LRTVLSLLPFA----LSGVLTPEISCRNEYGAVDWF-----IFYKLPRTSKASE 53  
 DB 40 LASVIFLLFAPPIVYFFIMACQYSCALT-GPVDIVTGHARLSDIWAFTPTIRKAAQ 98  
 QY 54 EAGL----QVLYLDSRQTWNKSL-----YLINSTRSALGRLOH 89  
 DB 99 LYTLVTFQVLLYTSLPDFCHKFLPGVVGIGQAGVTPAGVNVKYLQINGLQALLHLW 158  
 QY 90 LYDTHNSTNDTAYLIYNDGVPGSVNYSRQYGHAKGLLVNRTQGFWLHISVPKPPVHG 149  
 DB 159 FANAHLWSFPTIIPDNWIP-----LLWCAN--ILGYAVSTFAMVKGY 200  
 QY 150 EYPTSGRRYQGTGCTGTFGYSQFEEIDFQLLVLPNIYSCF-----IPSTFHWKLIY 201  
 DB 201 FFPTSARDCKFTG---NFFYNYVMGIEF---NPRIGKWFDFKLPFGNPGIVAWTLIN 252  
 QY 202 MPRMCANSSSLKTPVRYLAELHS--AQGLNFVHFSAKSFYTDITFTGWIQAQKLKTLAQ 259  
 DB 253 L-----SFAK-----QRELHSHVTVNAMLVNVLQ-AIYVIDFF--W-----NE 288  
 QY 260 TWOKKKQEL-----PSNCSLP-----Y 276  
 DB 289 TWYLTIDICHDFHGWYLGWDCVWLPYLYTLQGLVLYVHPVOLSTPHAVGVLLGLVGY 348  
 QY 277 HVYNI-----KSGVTSKSYFSS--QDHSKWCVS--IKGSANWTC 314  
 DB 349 YIFRVANHQKDLFRRTDGRCLIWGRPKVTECSYTSADGQRHSHKLLVSGFWGVARHFN 408  
 QY 315 IGLNLSLHQAALR-GGGFICTKNHYIOAF-----HKLYLYRYG 351  
 DB 409 VGDLMSLAYCLACGGHLLPYFYIYMAILLTHRCLEDRHCASKYG 456

Search completed: October 14, 2004, 00:28:06  
 Job time : 74.1983 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2004, 00:24:00 ; Search time 31.3671 Seconds  
(without alignments)  
582.636 Million cell updates/sec

Title: US-10-790-589-2

Perfect score: 1927

Sequence: 1 MTAKPLRTVLSLLFPALSGV.....KNHYIQAFHKILRYGFK 354

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*

1: /cgn2\_6/ptodata/2/1aa/5A COMB.pap:\*

2: /cgn2\_6/ptodata/2/1aa/5B COMB.pap:\*

3: /cgn2\_6/ptodata/2/1aa/6A COMB.pap:\*

4: /cgn2\_6/ptodata/2/1aa/6B COMB.pap:\*

5: /cgn2\_6/ptodata/2/1aa/PTUS COMB.pap:\*

6: /cgn2\_6/ptodata/2/1aa/backfiles1.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1927	100.0	354	4	US-09-574-942-2
2	1927	100.0	354	4	US-09-807-784B-1
3	1293.5	67.1	357	4	US-09-574-942-4
4	1286.5	66.8	361	4	US-09-807-784B-3
5	596.5	31.0	360	3	US-09-147-915-3
6	596.5	31.0	360	3	US-08-639-294-2
7	596.5	31.0	360	4	US-09-861-034B-2
8	443	23.0	275	3	US-09-147-915-4
9	236.5	12.3	192	4	US-09-210-422-2
10	101	5.2	481	4	US-09-328-352-7749
11	94	4.9	1385	1	US-07-675-280-2
12	94	4.9	1385	1	US-07-675-772-2
13	94	4.9	1385	1	US-08-063-170-2
14	94	4.9	1385	1	US-08-158-232-2
15	94	4.9	1385	1	US-08-304-626-2
16	94	4.9	1385	1	US-08-316-301A-2
17	94	4.9	1385	1	US-08-611-928-2
18	94	4.9	1385	3	US-09-173-891-2
19	94	4.9	1385	3	US-09-076-137-2
20	94	4.9	1385	4	US-09-738-363-2
21	94	4.9	1385	5	PCT-US92-03624-2
22	94	4.9	1385	6	5281530-1
23	94	4.9	1385	6	5426049-1
24	94	4.9	2165	1	US-08-514-975B-2
25	94	4.9	2165	5	PCT-US95-12507-2
26	93.5	4.9	583	6	5256558-4
27	90	4.7	1445	1	US-08-015-986A-2

28	90	4.7	1445	2	US-08-446-363-2	Sequence 2, Appli
29	89	4.6	267	1	US-08-015-973-3	Sequence 3, Appli
30	89	4.6	267	2	US-08-448-164-3	Sequence 3, Appli
31	89	4.6	267	3	US-08-081-929-3	Sequence 3, Appli
32	86.5	4.5	473	4	US-09-284-768A-24	Sequence 24, Appli
33	86.5	4.5	1198	4	US-09-284-768A-10	Sequence 10, Appli
34	85	4.4	587	4	US-09-328-352-8096	Sequence 8096, Ap
35	85	4.4	1442	1	US-08-015-986A-3	Sequence 3, Appli
36	85	4.4	1442	2	US-08-446-363-3	Sequence 3, Appli
37	84	4.4	675	4	US-08-426-509A-4	Sequence 4, Appli
38	84	4.4	675	4	US-08-232-545-4	Sequence 4, Appli
39	84	4.4	675	5	PCT-US95-05008-4	Sequence 4, Appli
40	83.5	4.3	335	4	US-09-134-001C-5021	Sequence 5021, Ap
41	82	4.3	393	4	US-09-107-532A-4397	Sequence 4397, Ap
42	82	4.3	520	3	US-08-810-003-5	Sequence 5, Appli
43	82	4.3	520	4	US-09-776-490-5	Sequence 5, Appli
44	82	4.3	1438	4	US-09-209-916-1	Sequence 1, Appli
45	82	4.3	1471	1	US-08-683-839B-3	Sequence 3, Appli

## ALIGNMENTS

### RESULT 1

US-09-574-942-2

; Sequence 2, Application US/09574942

; Patent No. 6358723

; GENERAL INFORMATION:

; APPLICANT: Eastman, Alan R.

; APPLICANT: Krieser, Ronald J.

; TITLE OF INVENTION: DEXYRIBONUCLEASE II BETA PROTEINS AND CDNAS

; FILE REFERENCE: DC-0137

; CURRENT APPLICATION NUMBER: US/09/574,942

; CURRENT FILING DATE: 2000-05-19

; PRIOR APPLICATION NUMBER: 09/541,840

; PRIOR FILING DATE: 2000-04-03

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; TYPE: PRT

; ORGANISM: Mus sp.

US-09-574-942-2

Query Match	100.0%;	Score 1927;	DB 4;	Length 354;
Best Local Similarity	100.0%;	Pred. No. 7.6e-206;		
Matches 354;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MTAKPLRTVLSLLFPALSGVLTPTPTSCNNEYGEADVDFIYKLPKRTSKASREAGLQYL	60	
Db	1	MTAKPLRTVLSLLFPALSGVLTPTPTSCNNEYGEADVDFIYKLPKRTSKASREAGLQYL	60	
QY	61	YLDSTRQWNKSLYLINSTRSALGRTLQHLXYDTHNSTNDTAYLIYNDGVPGSVNSRQYG	120	
Db	61	YLDSTRQWNKSLYLINSTRSALGRTLQHLXYDTHNSTNDTAYLIYNDGVPGSVNSRQYG	120	
QY	121	HAKGLLVNRTQCFWLIHSVPKFPVHGVEYPTSGRRYQGTGICITFGYSQFEEDFOLL	180	
Db	121	HAKGLLVNRTQCFWLIHSVPKFPVHGVEYPTSGRRYQGTGICITFGYSQFEEDFOLL	180	
QY	181	VLPQNIYSCFIPSTPHWKLIYMPRCANSSSLKIPVYLAELHSAQGLNFVHFAKSSFT	240	
Db	181	VLPQNIYSCFIPSTPHWKLIYMPRCANSSSLKIPVYLAELHSAQGLNFVHFAKSSFT	240	
QY	241	DDIFTGWIAQKLKTHLLAQTWQKKQELFNSCLSPHYVNIKSIQVTSKYSFSSRDHCK	300	
Db	241	DDIFTGWIAQKLKTHLLAQTWQKKQELFNSCLSPHYVNIKSIQVTSKYSFSSRDHCK	300	
QY	301	WCVSIKGSANRWTCIGDLNRSIHOALRGGGFTCTKNHYIQAFHKILRYGFCCK	354	
Db	301	WCVSIKGSANRWTCIGDLNRSIHOALRGGGFTCTKNHYIQAFHKILRYGFCCK	354	

## RESULT 2

US-09-807-784B-1  
; Sequence 1, Application US/09807784B  
; Patent No. 6653118  
; GENERAL INFORMATION:  
; APPLICANT: Tanuma, Sei-ichi  
; APPLICANT: Shiokawa, Daisuke  
; TITLE OF INVENTION: No. 6653118el Deoxyribonuclease, Gene Encoding Thereof and Use Thereof  
; FILE REFERENCE: 210792  
; CURRENT APPLICATION NUMBER: US/09/807,784B  
; CURRENT FILING DATE: 1999-08-17  
; PRIOR APPLICATION NUMBER: JP 11-230870  
; PRIOR FILING DATE: 1999-08-17  
; NUMBER OF SEQ ID NOS: 15  
; SEQ ID NO 1  
; LENGTH: 354  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-807-784B-1

Query Match 100.0%; Score 1927; DB 4; Length 354;  
Best Local Similarity 100.0%; Pred. No. 7.6e-206; Mismatches 0; Indels 0; Gaps 0;  
Matches 354; Conservative 0;

```
Qy 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYL 60
Db 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYL 60
Qy 61 YLDSTROTWNKSLYLINSTRSALGRTQLHLYDTHNS-TNDTAYLIYNDGVPGSVNSRYQY 120
Db 61 YLDSTROTWNKSLYLINSTRSALGRTQLHLYDTHNS-TNDTAYLIYNDGVPGSVNSRYQY 120
Qy 121 HAKGLLVNRTQGFVLIHSVPKPPV--HGYEYPTSGRRYGTGICITFGYSQFEIDFQLL 180
Db 121 HAKGLLVNRTQGFVLIHSVPKPPV--HGYEYPTSGRRYGTGICITFGYSQFEIDFQLL 180
Qy 181 VLQPNISCFIPSTFHWKLIYMPRMCMANSSSLKIPVRYLAELHSAOGLNFVHFAKSSFT 240
Db 181 VLQPNISCFIPSTFHWKLIYMPRMCMANSSSLKIPVRYLAELHSAOGLNFVHFAKSSFT 240
Qy 241 DDIFTGWIQAOKLTHLLAQTWQKKQELPNSCLPYHYVNIKSIQVTSKSYFSSRODHSK 300
Db 241 DDIFTGWIQAOKLTHLLAQTWQKKQELPNSCLPYHYVNIKSIQVTSKSYFSSRODHSK 300
Qy 301 WCVSIKGSANRWTCIGDLNLSHQAALRGGGFICTKNHVIYQAFHKLVLRYGFCK 354
Db 301 WCVSIKGSANRWTCIGDLNLSHQAALRGGGFICTKNHVIYQAFHKLVLRYGFCK 354
```

## RESULT 3

US-09-574-942-4  
; Sequence 4, Application US/09574942  
; Patent No. 6358723  
; GENERAL INFORMATION:  
; APPLICANT: Eastman, Alan R.  
; APPLICANT: Krieser, Ronald J.  
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS  
; FILE REFERENCE: DC-0137  
; CURRENT APPLICATION NUMBER: US/09/574,942  
; CURRENT FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: 09/541,840  
; PRIOR FILING DATE: 2000-04-03  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 357  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-574-942-4

Query Match 67.1%; Score 1293.5; DB 4; Length 357;  
Best Local Similarity 66.4%; Pred. No. 2.8e-135; Mismatches 45; Indels 3; Gaps 2;  
Matches 237; Conservative 45;

```
Qy 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYL 60
Db 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYL 60
Qy 61 YLDSTROTWNKSLYLINSTRSALGRTQLHLYDTHNS-TNDTAYLIYNDGVPGSVNSRYQY 119
Db 61 YLDSTRTSRWKSQQLMNDTKSVLGRTLQQLYEAVASKSNNTAYLIYNDGVPKPNYSRKY 120
Qy 120 CHAKGLLVNRTQGFVLIHSVPKPPV--HGYEYPTSGRRYGTGICITFGYSQFEIDF 177
Db 120 CHAKGLLVNRTQGFVLIHSVPKPPV--HGYEYPTSGRRYGTGICITFGYSQFEIDF 177
Qy 121 GHTKGLLVNRTQGFVLIHSVPKPPV--HGYEYPTSGRRYGTGICITFGYSQFEIDF 180
Db 121 GHTKGLLVNRTQGFVLIHSVPKPPV--HGYEYPTSGRRYGTGICITFGYSQFEIDF 180
Qy 178 QLLVLPQNIYSCFIPSTFHWKLIYMPRMCMANSSSLKIPVRYLAELHSAOGLNFVHFAKSS 237
Db 178 QLLVLPQNIYSCFIPSTFHWKLIYMPRMCMANSSSLKIPVRYLAELHSAOGLNFVHFAKSS 237
Qy 181 QLLVLPQNIYSCFIPSTFHWKLIYMPRMCMANSSSLKIPVRYLAELHSAOGLNFVHFAKSS 240
Db 181 QLLVLPQNIYSCFIPSTFHWKLIYMPRMCMANSSSLKIPVRYLAELHSAOGLNFVHFAKSS 240
Qy 238 FYTDDIFTGWIQAOKLTHLLAQTWQKKQELPNSCLPYHYVNIKSIQVTSKSYFSSROD 297
Db 238 FYTDDIFTGWIQAOKLTHLLAQTWQKKQELPNSCLPYHYVNIKSIQVTSKSYFSSROD 297
Qy 241 SFLDGIFAAWQAOKLTHLLAQTWQKKQELPNSCLPYHYVNIKSIQVTSKSYFSSROD 300
Db 241 SFLDGIFAAWQAOKLTHLLAQTWQKKQELPNSCLPYHYVNIKSIQVTSKSYFSSROD 300
Qy 298 HSKWCYSIKGSANRWTCIGDLNLSHQAALRGGGFICTKNHVIYQAFHKLVLRYGFCK 354
Db 298 HSKWCYSIKGSANRWTCIGDLNLSHQAALRGGGFICTKNHVIYQAFHKLVLRYGFCK 354
Qy 301 HAKWCISQKGTKNRWTCIGDLNLSHQAALRGGGFICTKNHVIYQAFHKLVLRYGFCK 357
Db 301 HAKWCISQKGTKNRWTCIGDLNLSHQAALRGGGFICTKNHVIYQAFHKLVLRYGFCK 357
```

## RESULT 4

US-09-807-784B-3  
; Sequence 3, Application US/09807784B  
; Patent No. 6653118  
; GENERAL INFORMATION:  
; APPLICANT: Tanuma, Sei-ichi  
; APPLICANT: Shiokawa, Daisuke  
; TITLE OF INVENTION: No. 6653118el Deoxyribonuclease, Gene Encoding Thereof and Use Thereof  
; FILE REFERENCE: 210792  
; CURRENT APPLICATION NUMBER: US/09/807,784B  
; CURRENT FILING DATE: 1999-08-17  
; PRIOR APPLICATION NUMBER: JP 11-230870  
; PRIOR FILING DATE: 1999-08-17  
; NUMBER OF SEQ ID NOS: 15  
; SEQ ID NO 3  
; LENGTH: 361  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-807-784B-3

Query Match 66.8%; Score 1286.5; DB 4; Length 361;  
Best Local Similarity 66.1%; Pred. No. 1.7e-134; Mismatches 45; Indels 3; Gaps 2;  
Matches 236; Conservative 45;

```
Qy 1 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYL 60
Db 5 MTAKPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASEAGLQYL 64
Qy 61 YLDSTROTWNKSLYLINSTRSALGRTQLHLYDTHNS-TNDTAYLIYNDGVPGSVNSRYQY 119
Db 65 YLDSTRTSRWKSQQLMNDTKSVLGRTLQQLYEAVASKSNNTAYLIYNDGVPKPNYSRKY 124
Qy 120 CHAKGLLVNRTQGFVLIHSVPKPPV--HGYEYPTSGRRYGTGICITFGYSQFEIDF 177
Db 125 GHTKGLLVNRTQGFVLIHSVPKPPV--HGYEYPTSGRRYGTGICITFGYSQFEIDF 184
Qy 178 QLLVLPQNIYSCFIPSTFHWKLIYMPRMCMANSSSLKIPVRYLAELHSAOGLNFVHFAKSS 237
Db 185 QLLVLPQNIYSCFIPSTFHWKLIYMPRMCMANSSSLKIPVRYLAELHSAOGLNFVHFAKSS 244
Qy 238 FYTDDIFTGWIQAOKLTHLLAQTWQKKQELPNSCLPYHYVNIKSIQVTSKSYFSSROD 297
Db 245 SFLDGIFAAWQAOKLTHLLAQTWQKKQELPNSCLPYHYVNIKSIQVTSKSYFSSROD 304
Qy 298 HSKWCYSIKGSANRWTCIGDLNLSHQAALRGGGFICTKNHVIYQAFHKLVLRYGFCK 354
Db 305 HAKWCISQKGTKNRWTCIGDLNLSHQAALRGGGFICTKNHVIYQAFHKLVLRYGFCK 361
```



```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/639,294
; FILING DATE: 25-Apr-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Johnston, Sean A.
; REGISTRATION NUMBER: 35,910
; REFERENCE/DOCKET NUMBER: P1024
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3562
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
;
US-08-639-294-2

Query Match 31.0%; Score 596.5; DB 3; Length 360;
Best Local Similarity 38.0%; Pred. No. 1.2e-57;
Matches 132; Conservative 56; Mismatches 144; Indels 15; Gaps 7;

QY 17 LSGVLGTP--ETISCRNVEGEAVDWFYFKLPK-RTSKASEAGLQVLYLDSTRQTWNKSL 73
DB 6 LAALLCVPALGATCYGDSQPDVDFVYKLPALRGSGEAAQRLQVLYLDSSGGWRDGR 65
QY 74 YLIISTRSALGRTHQLHYDTHSNTPAYLIYND--GVPGSVNYSRQYGHAKGLLVNRT 131
DB 66 ALINSPEGAVERSLQPLY--RSNTSLAFLLYNDQPPQSKAQDSSMRGHTKGVLLDHD 123
QY 132 QGFWLHISVPKPPP---VHCYEYPTSGRYGTCITFTGYSOFEIDTQLLVLPNIYS 188
DB 124 GGFWLHVSVPNPPPPPASSAAYSWFHSACTYGTQLLCSFPFQAQFSKMGKQLTYTPWVYN 183
QY 189 CRIPSTFHWKLIYMPMCANSSSLKIPVRYLAELHSAQGLINVFVHFAKSGFYTDITGWI 248
DB 184 YQLEGIFAQEPFDLENVVGHHVSQEPWNSSITLTSGAGAVTQSFATKFKGDDLYSGWL 243
QY 249 AQKLKTHLAQWQKKKQBLPNCSLPYHYNIKISGVTSKS--YFSSQDHSKWCVSIK 306
DB 244 AAALGNLQVQVHKTVTGTLPSNCSDSIQWLVNINQIAFPGPAGSPFNSTEDHSKWCVSPK 303
QY 307 GSANRWTCIGDLNRSLHQALRGGGGCTCKNHVITYQAFHKLILRYGFC 353
DB 304 GP---WTCVGDNRNGEHRQGGGTLCAQLPALWKAFOPLVKNYQFC 347

RESULT 7
US-09-861-034B-2
; Sequence 2, Application US/09861034B
; Patent No. 6569429
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; Baron, Will F.
; TITLE OF INVENTION: Human DNase II
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/861 034B

```

```

; FILING DATE: 18-May-2001
; CLASSIFICATION: <Unknown>
; PRIORITY INFORMATION:
; APPLICATION NUMBER: 08/639294
; FILING DATE: 25-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Johnston, Sean A.
; REGISTRATION NUMBER: 35,910
; REFERENCE/DOCKET NUMBER: F1024D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3562
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-861-034B-2

```

```

Query Match 31.0%; Score 596.5; DB 4; Length 360;
Best Local Similarity 38.0%; Pred. No. 1.2e-57;
Matches 132; Conservative 56; Mismatches 144; Indels 15; Gaps 7;

QY 17 LSGVLGTP--EISCRNEYGAVDWFIYKLPK-RTSKASEAGLQYLYLDSTRQTNKSL 73
Db 6 LAALLCPAGALICYGDSGGPWFVYKLPALRGSGEAAQGLQYLYLDESSGWRDGR 65
QY 74 YLINSTRSALGRTOHLYDTHNSTNDYLYND--GVPGSVNYSROYGHAKGLLVNRT 131
Db 66 ALINSPGAVGRSLQPLY--RSNTSLQALFLYNDQPPQPSKAQDSMRGHTKGLVLDHD 123
QY 132 QGFWLHVSVPKPP--VHGVEYPTSGRRYCGTICITFCYSOFEEIDFOLLVLQNIYS 188
Db 124 GGFWLHVSVPKPP--VHGVEYPTSGRRYCGTICITFCYSOFEEIDFOLLVLQNIYS 183
QY 189 CFTPTSPFHMKLIYMPRCANSSSLKIPVRYLAELHSAQGLNFVHFVFAKSSFYTDIFGTWI 248
Db 184 YQLEGFAQEPDLENVVGKHHVSOEPWNSITLTSQAGAVFQSFQKFGDLSGWL 243
QY 249 AQKLTLLAQTQWKKQELPNSCLPYHYVNIKSIGVTSKS--YFSSRODHKWCVSIG 306
Db 244 AAALGTNLQVQFHKTVGILPNSCSDIWLQVLNVNQLAFPGPAGPSPFNSTEDHSHKWCVSPK 303
QY 307 GSNRWTCIGDLNRLHQAARGGFGICTKKNHYLYQAFHKLKLYRGFC 353
Db 304 GP---WTCVGDMMNRNOGEORGGTLCALPALWKAQPLVKNYQPC 347

```

```

RESULT 8
US-09-147-915-4
; Sequence 4, Application US/09147915A
; Patent No. 6184034
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan
; APPLICANT: Krieser, Ronald
; TITLE OF INVENTION: Deoxyribonuclease II Proteins and cDNAs
; FILE REFERENCE: DC-0097
; CURRENT APPLICATION NUMBER: US/09/147,915A
; CURRENT FILING DATE: 1999-03-23
; EARLIER APPLICATION NUMBER: PCT/US97/18262
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/028,539
; EARLIER FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Bos sp.
US-09-147-915-4

```

Query Match 23.0%; Score 443; DB 3; Length 275;

```

Best Local Similarity 36.0%; Pred. No. 9.6e-41;
Matches 94; Conservative 47; Mismatches 108; Indels 12; Gaps 5;

QY 101 AXLIYNDGVPGSVNY--SRQYGHAKGLLVNRTQGFWLHVSVPKPP--VHGVEYPTSG 155
Db 4 AFVLYNDQPPKSSSKDSSSRGHTKGLVLLDQGGFWLHVSVPKPPRASSAAISWPPGA 63
QY 156 RRYGTCGICITFCYSOFEEIDFOLLVLQNIYSCFIPSTFHWKLIYMPRCANSSSLKIP 215
Db 64 QKYGQTLICUSPPLTOFLDISKQLTYYPLVYDHRLEGFGQKFFYLEEVVKGHHVROGP 123
QY 216 VRYLAELHSAQGLNFVHFVFAKSSFYTDIFGTWIAQKLKTHLLAQTQWKKQELPNSCLP 275
Db 124 WNSSVLTSTKKGATFQSFQKFGDLYSGWLAEALGSTLQVQFQWQSSGILPNSCSGA 183
QY 276 YHYVNIKSIGVTSKS--YFSSRODHSHKWCVSIGKSNRWTCIGDLNRLHQAARGGFGIC 333
Db 184 QHVFDTQTAFPGPAGPAGFATNEDHSHKWCVTPKGP---WACVGDMMNRNOGEORGGTLC 240
QY 334 TKNHYYIQAFHKLKLYRGFC 354
Db 241 AQ--MLWKAFKPLVKAWEPC 259

```

```

RESULT 9
US-09-210-422-2
; Sequence 2, Application US/09210422A
; Patent No. 6455250
; GENERAL INFORMATION:
; APPLICANT: AGUILERA, RENATO J.
; APPLICANT: LYON, CHRISTOPHER J.
; TITLE OF INVENTION: ENDONUCLEASE COMPOSITIONS AND METHODS OF USE
; FILE REFERENCE: UCLA019
; CURRENT APPLICATION NUMBER: US/09/210,422A
; CURRENT FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 60/069,205
; PRIOR FILING DATE: 1997-12-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Musca sp.
US-09-210-422-2

```

```

Query Match 12.3%; Score 236.5; DB 4; Length 192;
Best Local Similarity 27.1%; Pred. No. 5.4e-18;
Matches 67; Conservative 27; Mismatches 96; Indels 57; Gaps 4;

QY 6 LRTVLGILFPALGVLGTPETISCRNEYGAVDWFIYKLPKRTSKASEAGLQYLYLDST 65
Db 1 MATYRSLLLAALLWV--PAEALSCVGSQGPVDFVYVVKLFAHSGSRDTPKGLTCKYWDQ 59
QY 66 RQTNKSLYLINSTRSALGRTOHLYDTHNSTNDYLYNDGV--GSVNYSGROYGHAK 123
Db 60 SDGWQDGVGYNSPEGAVGRSLQPLY--RKNSSQLAFLYNDQPPKSSSTRDSTGHGHTK 117
QY 124 GLLVNRTQGFWLHVSVPKPPVHGVEYPTSGRRYCGTICITFCYSOFEEIDFOLLVLQ 183
Db 118 GK-----QLITYTY 125
QY 184 PNIYSCFIPSTFHWKLIYMPRCANSSSLKIPVRYLAELHSAQGLNFVHFVFAKSSFYTDI 243
Db 126 PLVYDHKLEGFFAQKLPDLTETVKNQHLHPEWNSVILTQAGATFQSFQKFGKFGDDL 185
QY 244 FTGWIQA 250
Db 186 YSGWLAE 192

```

```

RESULT 10
US-09-328-352-7749
; Sequence 7749, Application US/09328352

```

```
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7749
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-7749

Query Match 5.2%; Score 101; DB 4; Length 481;
Best Local Similarity 23.8%; Pred. No. 0.029;
Matches 55; Conservative 27; Mismatches 83; Indels 72; Gaps 12;

QY 139 SVKFPFVHGVEYPTSGRRYGGTGCITFGYSQFEEDFOLLVLPNIYSCPTPS-----193
Db 234 SQAKSPTEHG---SLAKDYAQOHPVVKAYLQATIEANFLIQEQPEKYSFPAETGIP 290
QY 194 -----TFH-----WKLIYPRMCANSSSLK-----IPVRYLAELH 223
Db 291 AEVVYLFHGPLGLQTRDLTWKPEYRKATQIAIDTLKVLGKNDGCTLDVKNKFDIDQYIKDAF 350
QY 224 SAQGLNF---VHFAPKSSFTDDIFTGWTAAQKLT-HLLAQTWKKKQELPNSCSIPYHV 278
Db 351 QASGLNYSQALADYAKSPVANDALG---QPIKTFDQVTVQIIVKGEKVRSS-YETPEHA 406
QY 279 YN-----IKSIGVTSKYSFSSRODHSK-----WCVS-----IKGSANRW 312
Db 407 FSDLKKIQAGKTVRVVYS--QDHQSDKILLANLAWYATDKAGIQIAFLIKDDAEKW 461

RESULT 11
US-07-876-280-2
; Sequence 2, Application US/07876280
; Patent No. 5262158
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel M.
; APPLICANT: Cannon, Raymond J.C.
; APPLICANT: Bagley, Angela L.
; TITLE OF INVENTION: No. 5262158el Bacillus thuringiensis Isolates for
; TITLE OF INVENTION: Controlling Acarides
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/876,280
; FILING DATE: 19920430
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: M/S 104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1385 amino acids
```

```
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: BACILLUS THURINGIENSIS
; INDIVIDUAL ISOLATE: PSI7
; IMMEDIATE SOURCE:
; CLONE: E. coli NM522(pMYC 1627) NRRL B-18651
; US-07-876-280-2

Query Match 4.9%; Score 94; DB 1; Length 1385;
Best Local Similarity 20.5%; Pred. No. 0.97;
Matches 76; Conservative 48; Mismatches 134; Indels 112; Gaps 16;

QY 31 EYGEADVDFIFKLPKRTSKASEAGLQVLYLDSTR-----QTWNK-----71
Db 399 EVDSDVETRLYGQLP---AVDPQAGNIVSDISSNPIIQINMDTWKTPPQAGSGWNTNL 454
QY 72 -----SLYLINSTRSALGR-----TLQHLDYDTHNSTNDTAYLIYNDGVP--GSVNY 115
Db 455 MRGSVSGLSFLQKRDGTRLSAGMGGGFADTIYSLPATH-----YLSYLYGTPYQYSDNY 507
QY 116 SRQYGHAKGLLVNRTQGFWLHSPVK---FPPVHGVEYPTSGRRYGGTGCITFGYSQF 172
Db 508 SHVGLAVGV-----STPQEAATLENIIG-----QPDEQCNVSTMGF-PF 545
QY 173 EEDFDQLLVLPQNIYSCFIPSTFHWKLIYPRMCANSSSLKIPVRYLAELHS-----224
Db 546 EKASYGGTVVKEWLGANAMKLSPGQSIGIPIPTNVTSGEYQIRCVASNDNTNVFNVD 605
QY 225 -----AQGLNFVHFAPKSSFTDDIFTGWTAAQKLTLLAQ-----259
Db 606 GGANPIFQIQINEASTVDNNTGVQGGANGVYVVKSIATDNTSFTPEIPAKTINVHLTNQSSD 665
QY 260 TWQKKKQELPNSCSIPYH-VYNIKIGVTSKYSFSSRODHSKVCVSIKGSANRWTCIGDL 318
Db 666 VELDRIEFIPFSLPLTYHGSYNTSS-GADDVLWSSSNMNY--YDIIVNGQANS-----SSI 718
QY 319 NRSLSHQALRG 328
Db 719 ASSMELLNKG 728

RESULT 12
US-07-675-772-2
; Sequence 2, Application US/07675772
; Patent No. 5262399
; GENERAL INFORMATION:
; APPLICANT: Hickie, Leslie A.
; APPLICANT: Sick, August J.
; APPLICANT: Schwab, George E.
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Payne, Jewel M.
; TITLE OF INVENTION: No. 5262399el Compositions and Methods for the Control of
; TITLE OF INVENTION: Flukes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROMAN SALIWANCHIK
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/675,772
```

FILING DATE: 19910327  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: SALIWANCHIK, ROMAN  
REGISTRATION NUMBER: 21,023  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 904-375-8100  
TELEFAX: 904-372-5800  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1385 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: BACILLUS THURINGIENSIS  
INDIVIDUAL ISOLATE: PS17  
IMMEDIATE SOURCE:  
LIBRARY: LAMBDA GEM (TM) - 11 LIBRARY OF KENNETH NARVA  
CLONE: 17A  
US-07-675-772-2

Query Match 4.9%; Score 94; DB 1; Length 1385;  
Best Local Similarity 20.5%; Pred. No. 0.97; 134; Indels 112; Gaps 16;  
Matches 76; Conservative 48; Mismatches 48; Indels 112; Gaps 16;

QY 31 EYGEAVDWFIFYKLPKRTSKASEAGLYLYLDSTR-----QTNK----- 71  
DB 399 EVQDSVETRLYGQLP-----AVDPOAGPNYVSIDSSNPIQINMDTWKTPPGASGWNINL 454  
QY 72 -----SLYLINSTSALGR-----TLQHLVDTHNSTNDTAYLYNDGVP--GSVNY 115  
DB 455 MRGSVGLSFLQRDGTLSAGMGGFADTIYSLPATH-----YLSLYGTPYQTSNDY 507  
QY 116 SRQYGHAKGLLVNRTQGFWLHISVPK---PPPVHGYEYPTSGRRYGTGICITFGYSQF 172  
DB 508 SGHVGLVGV-----STQEAFLPNIIG-----QPDQGNVSTMGP-PF 545  
QY 173 BEIDFQLLVLPQNIYSCFIPSTFWKLIYMPRCANSSSLKIPVRYLAELHS----- 224  
DB 546 EKASYGTVVKEWLNAGANAMKLSFGSIGIPITNVTSGEYQIRCRYASNDNTNFFENVDT 605  
QY 225 -----AQGLNFVHFAKSSFYDDIFTGWIQAOKLKTLLAQ----- 259  
DB 606 GGANPIFQINFASTVDNNTGVQANGVYVVKYSATTDSNFTPAKTINVHLTNQSSD 665  
QY 260 TWQKKQELPNSCLPYH-VYNIKSIGVTSKYSFSSRQDHSKWCVKSGSANRWTCIGDL 318  
DB 666 VLDRIEIPFSLPLIYHGSYNTSS-GADDVYLVSSNNY--YDIIVNGQANS-----SSI 718  
QY 319 NRSLHQALRG 328  
DB 719 ASSMHLNKG 728

RESULT 13  
US-08-063-170-2  
Sequence 2, Application US/08063170  
Patent No. 5350576  
GENERAL INFORMATION:  
APPLICANT: Kim, Leo  
APPLICANT: Schwab, George E.  
TITLE OF INVENTION: Compositions and Methods for Inducing an Immune  
TITLE OF INVENTION: Response for Protection Against Endoparasites and Exoparasites  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DAVID R. SALIWANCHIK  
STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: FL

COUNTRY: USA  
ZIP: 32606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/063,170  
FILING DATE: 19930517  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,141  
FILING DATE: 30-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/759,248  
FILING DATE: 13-SEP-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: SALIWANCHIK, DAVID R.  
REGISTRATION NUMBER: 31,794  
REFERENCE/DOCKET NUMBER: M/S 103.C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 904-375-8100  
TELEFAX: 904-372-5800  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1385 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Bacillus thuringiensis  
STRAIN: PS17  
INDIVIDUAL ISOLATE: PS17a  
US-08-063-170-2

Query Match 4.9%; Score 94; DB 1; Length 1385;  
Best Local Similarity 20.5%; Pred. No. 0.97; 134; Indels 112; Gaps 16;  
Matches 76; Conservative 48; Mismatches 48; Indels 112; Gaps 16;

QY 31 EYGEAVDWFIFYKLPKRTSKASEAGLYLYLDSTR-----QTNK----- 71  
DB 399 EVQDSVETRLYGQLP-----AVDPOAGPNYVSIDSSNPIQINMDTWKTPPGASGWNINL 454  
QY 72 -----SLYLINSTSALGR-----TLQHLVDTHNSTNDTAYLYNDGVP--GSVNY 115  
DB 455 MRGSVGLSFLQRDGTLSAGMGGFADTIYSLPATH-----YLSLYGTPYQTSNDY 507  
QY 116 SRQYGHAKGLLVNRTQGFWLHISVPK---PPPVHGYEYPTSGRRYGTGICITFGYSQF 172  
DB 508 SGHVGLVGV-----STQEAFLPNIIG-----QPDQGNVSTMGP-PF 545  
QY 173 BEIDFQLLVLPQNIYSCFIPSTFWKLIYMPRCANSSSLKIPVRYLAELHS----- 224  
DB 546 EKASYGTVVKEWLNAGANAMKLSFGSIGIPITNVTSGEYQIRCRYASNDNTNFFENVDT 605  
QY 225 -----AQGLNFVHFAKSSFYDDIFTGWIQAOKLKTLLAQ----- 259  
DB 606 GGANPIFQINFASTVDNNTGVQANGVYVVKYSATTDSNFTPAKTINVHLTNQSSD 665  
QY 260 TWQKKQELPNSCLPYH-VYNIKSIGVTSKYSFSSRQDHSKWCVKSGSANRWTCIGDL 318  
DB 666 VLDRIEIPFSLPLIYHGSYNTSS-GADDVYLVSSNNY--YDIIVNGQANS-----SSI 718  
QY 319 NRSLHQALRG 328  
DB 719 ASSMHLNKG 728

RESULT 14

US-08-158-232-2  
; Sequence 2, Application US/08158232  
; Patent No. 5596071  
; GENERAL INFORMATION:  
; APPLICANT: Payne, Jewel  
; APPLICANT: Kennedy, M. Keith  
; APPLICANT: Randall, John Brooks  
; APPLICANT: Meier, Henry  
; APPLICANT: Uick, Heidi Jane  
; APPLICANT: Poncetrada, Luis  
; APPLICANT: Schnepf, H. Ernest  
; APPLICANT: Schwab, George E.  
; APPLICANT: Fu, Jenny  
; TITLE OF INVENTION: No. 5596071el Bacillus thuringiensis Toxins Active  
; TITLE OF INVENTION: Against Hymenopteran Pests  
; NUMBER OF SEQUENCES: 51  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David R. Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: USA  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/158,232  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/887,980  
; FILING DATE: 22-MAY-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/797,645  
; FILING DATE: 25-NOV-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/703,977  
; FILING DATE: 22-MAY-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saliwanchik, David R.  
; REGISTRATION NUMBER: 31,794  
; REFERENCE/DOCKET NUMBER: M/SCJ104.C1  
; TELEPHONE: 904-375-8100  
; TELEFAX: 904-372-5800  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1385 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; STRANDEDNESS: single  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: YES  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: BACILLUS THURINGIENSIS  
; STRAIN: PSI7  
; INDIVIDUAL ISOLATE: PSI7a  
; IMMEDIATE SOURCE:  
; CLONE: E. coli NM522 (pMYC1627) NRRL B-18651  
US-08-158-232-2  
Query Match 4.9%; Score 94; DB 1; Length 1385;  
Best Local Similarity 20.5%; Pred. No. 0.97;  
Matches 76; Conservative 48; Mismatches 134; Indels 112; Gaps 16;  
31 EYGRAVDWFIYKLPKRTSKASEAGLOVLYLDSTR-----QTWNK----- 71  
399 EVQDSVETLYGQLP-----AVDPQGPNYVSLDSNPFIQINMDTWKTPQAGSWNTNL 454

QY 72 -----SLYLINSTRSALGR-----TIQHLVDTNSTNDTAYLIYNDGVP--GSVNY 115  
Db 455 MRGSVGLSFLQRDGTLSAGMGCGFADTIYSLPATH-----YLSYLYGTPTQTSQNY 507  
QY 116 SROVGHAKGLLVNRTQGFWLIHSVPK---FPPVHGYEYPTSGRRYGTGTGICITFGYSQF 172  
Db 508 SCHVGALGV-----STPQEATLPIIIG-----QPDQGNVSTWGF-PF 545  
QY 173 EBTDFQLLVQPNYISCFIPSTFWKLIYMRMCANSSSLKIPVRYLAELHS----- 224  
Db 546 EKASYGTVVKEWLVGANAMKLSFGQSIGIPITNVTSGEYQIRCRYASNDNTNVPFVNDT 605  
QY 225 -----AQLNFVHFAKSFYTDITFTGWIAQKLTLLIAQ----- 259  
Db 606 GGANPIFQQINFASVTVDNNTGVQANGVYVYKSIATTDNSFTETPAKTINVHLTNQSSD 665  
QY 260 TWOKTKQELPSNGSLPYH-VYNIKSIIGVTSKVSFSSRQDHSKWCVSIKGSANRWTCIGDL 318  
Db 666 VELDRIEFIPFSLPIYHGSYNTSS-GADDVLWSSSNMNY--YDIIVNGOANS-----SSI 718  
QY 319 NRSLSHQAURG 328  
Db 719 ASSMHLINKG 728  
RESULT 15  
US-08-304-626-2  
; Sequence 2, Application US/08304626  
; Patent No. 5616495  
; GENERAL INFORMATION:  
; APPLICANT: Payne, Jewel M.  
; APPLICANT: Kennedy, M. Keith  
; APPLICANT: Randall, John Brooks  
; APPLICANT: Meier, Henry  
; APPLICANT: Uick, Heidi Jane  
; APPLICANT: Poncetrada, Luis  
; APPLICANT: Schnepf, Harry B.  
; APPLICANT: Schwab, George E.  
; TITLE OF INVENTION: No. 5616495el Bacillus thuringiensis Isolates  
; TITLE OF INVENTION: Active Against Hymenopteran Pests and Genes Encoding  
; NUMBER OF SEQUENCES: 39  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David R. Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: USA  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/304,626  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/887,980  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saliwanchik, David R.  
; REGISTRATION NUMBER: 31,794  
; REFERENCE/DOCKET NUMBER: M/SCJ 104  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 904-375-8100  
; TELEFAX: 904-372-5800  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1385 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: BACILLUS THURINGIENSIS  
STRAIN: PS17  
INDIVIDUAL ISOLATE: PS17a  
IMMEDIATE SOURCE:  
CLONE: E. coli NM522 (pMYC1627) NRRL B-18651  
US-08-304-626-2

Query Match 4.9%; Score 94; DB 1; Length 1385;  
Best Local Similarity 20.5%; Pred. No. 0.97;  
Matches 76; Conservative 48; Mismatches 134; Indels 112; Gaps 16;

Qy	31	EYGEAVDWTFYFKLPKRTSKASBEAGLQYLYLDSTR-----QTWNK-----	71
Db	399	EVQDSVETRLYGQLP-----AVDPQAGENYYSIDSSNFIQINMDTWKTPPQGASGWNINL	454
Qy	72	-----SYLINSTRSALGR-----TLQHLVDTHNSTNDTAYLIYNDGVP--GSVNY	115
Db	455	MRGSVSGLSFLQDGTRLSAGMGGFADTIYSLPATH-----YLSYLYGTPYQTSQNY	507
Qy	116	SRQYGHAKGLLVNRTQGFLLHSPK---PPVHGYEYPTSGRRYGTGICITFGYSQF	172
Db	508	SGHVGLVGV-----STPORATLPLNIG-----QPDEQGNVSTMGP-PF	545
Qy	173	HEIDFQLLVLPNTYSCFIPSTFHWKLIYMPRCANSSSLKIPVRYLAELHS-----	224
Db	546	EKASYGGTVVKEWLNGANAMKLSFGQSIGIPITNVTSGEYQRCRYASNDNTNVFENVDT	605
Qy	225	-----AQLNFVHFAKSSFYDDIFTGWIQKLTLLAQ-----	259
Db	606	GGANPIFQIINFASVDNNTGVQANGVYVVKSIATDINSFTEIPAKTINVHLTNQSSD	665
Qy	260	TWQKKQELPSNCSLPYH-VYNIKISIGVTSKSYFSSRDHSKWCVKISKGSANRWTCIGDL	318
Db	666	VFLDRIBFIPFSLPLIVHGSINTSS-GADDVWSSNNY--YDIIVNGQANS-----SSI	718
Qy	319	NRSLHQLRG 328	
Db	719	ASSMHLNKG 728	

Search completed: October 14, 2004, 00:33:35  
Job time : 33.3671 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2004, 00:41:52 ; Search time 2984.85 Seconds  
(without alignments)  
12245.607 Million cell updates/sec

Title: US-10-790-589-1

Perfect score: 1224

Sequence: 1 tccagtcctccatggaa.....aacgtagctccaataaaag 1224

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_hc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_hc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pin:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1216.6	99.4	2689	11 AK079224	AK079224 Mus muscu
2	595.4	48.6	695	14 CF725003	CF725003 UI-M-G20-
3	510	41.7	510	9 A1526873	A1526873 uj47d10.x
4	485.2	39.6	893	13 BQ681950	BQ681950 AGENCOURT

## RESULT 1

AK079224

LOCUS

DEFINITION

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224

AK079224





902 CCAAGTCTTACTTCACTTCTCGCAAGACCATTCCTCAATGCTGTGTTTCCATAAAGGGCT 961  
 Db CCAAGTCTTACTTCACTTCTCGCAAGACCATTCCTCAATGCTGTGTTTCCATAAAGGGCT 1256  
 962 CCGCAATCGCTGAGCTGCTGAGAGACCTAATCGAAGCCCTACACCAAGCCTTAAGAG 1021  
 Db CCGCAATCGCTGAGCTGCTGAGAGACCTAATCGAAGCCCTACACCAAGCCTTAAGAG 1316  
 1022 GTGGAGGATTCATCTGACAAAGATCACTATATACAGGATTCATTAATATATC 1081  
 Db GTGGAGGATTCATCTGACAAAGATCACTATATACAGGATTCATTAATATATC 1376  
 1082 TCCGTTATGGTCTCTGTAAGTAACTCGGTGAAGCCACACCTCTGTCCTTGAAGACA 1141  
 Db TCCGTTATGGTCTCTGTAAGTAACTCGGTGAAGCCACACCTCTGTCCTTGAAGACA 1436  
 1142 GTGGACATGGAACATCTGCTGAGTCTGTTCTCATTAATATCAAGGCTTCGAGTGAG 1201  
 Db GTGGACATGGAACATCTGCTGAGTCTGTTCTCATTAATATCAAGGCTTCGAGTGAG 1496  
 1202 CACAACGTAGCTCCCAATAAAG 1224  
 Db CACAACGTAGCTCCCAATAAAG 1519

RESULT 2  
 CF725003  
 LOCUS  
 DEFINITION UI-M-G20-cjo-i-11-0-UI.r1 NIH\_BMAP\_G20 Mus musculus cDNA clone  
 IMAGE:30606058 5', mRNA sequence.  
 CF725003  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus (house mouse)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NIH-MGC <http://mgi.nhl.nih.gov/>  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: Dr. James Lin University of Iowa  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Distribution information can be found at  
<http://genome.uiowa.edu/distribution/mousefl.html>  
 This clone was contributed by the Brain Molecular Anatomy Project  
 (BMAP)

Seq primer: pYX-5.  
 Location/Qualifiers  
 1. .695  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:30606058"  
 /tissue\_type="whole eye"  
 /dev\_stage="embryo 12.5,13.5,14.5 dpc"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_lib="NIH\_BMAP\_G20"  
 /note="Organ: Eye; Vector: pYX-Asc; Site 1: EcoR I;  
 Site 2: Not I; The library was constructed according  
 Bonaldi, Lennon and Soares, Genome Research, 6:791-806,  
 1996. Denatured RNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with oligo-dT  
 primer containing a Not I site. Double strand cDNA was  
 size selected according to Not I site. Double strand cDNA was  
 with EcoR I adaptor, digested with NotI and then cloned  
 directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail  
 is TTAATGAAGT. This library was created for the University  
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the  
 Developing Mouse Nervous System', supported by National  
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,  
 program coordinator."

ORIGIN

Query Match 48.6%; Score 595.4; DB 14; Length 695;  
 Best Local Similarity 99.2%; Pred. No. 2.3e-175;  
 Matches 640; Conservative 0; Mismatches 1; Indels 4; Gaps 2;  
 QY 2 CCAGTCCCTCATGGAATGAAGCCACAGATAGAAATGACAGCAAGCCCTTAAGAA 61  
 Db CACAGTCCCTCATGGAATGAAGCCACAGATAGAAATGACAGCAAGCCCTTAAGAA 111  
 QY 62 CAGTCTCTTTTGTCTCTTTTGGGCTCTCTGGGCTCTGGGACACCAAGAAATCTCAT 121  
 Db CAGTCTCTTTTGTCTCTTTTGGGCTCTCTGGGCTCTGGGACACCAAGAAATCTCAT 171  
 QY 122 GCAGAAATGAATATGATGATGAGCTGTGAGCTGTGATCTTTTATTAAGTTACCCAAAGGA 181  
 Db GCAGAAATGAATATGATGATGAGCTGTGAGCTGTGATCTTTTATTAAGTTACCCAAAGGA 231  
 QY 182 CTAGCAAGGCAAGTGAAGAGGGGGCTGACAGTACCTCTGACCTGACCTCCCAAGACAA 241  
 Db CTAGCAAGGCAAGTGAAGAGGGGGCTGACAGTACCTCTGACCTCCCAAGACAA 291  
 QY 242 COTGGAACAAGAGCCTCTACCTGATTAACAGCAGGAGTCTCTGGGAGGACCTTAC 301  
 Db COTGGAACAAGAGCCTCTACCTGATTAACAGCAGGAGTCTCTGGGAGGACCTTAC 351  
 QY 302 AGCATCTGTATGACACACATTAATCCAGATGACACAGCTTATTAATATACAGATG 361  
 Db AGCATCTGTATGACACACATTAATCCAGATGACACAGCTTATTAATATACAGATG 411  
 QY 362 GTGTCCCTGGATCTGTGAATTAACAGCAGATGACATGCCCAAGGTCCTGGTAT 421  
 Db GTGTCCCTGGATCTGTGAATTAACAGCAGATGACATGCCCAAGGTCCTGGTAT 471  
 QY 422 GGAACAGAACCCAGGGGTTCTGGCTGTATACACTCTGTTCCTCAAGTTTCCCAAGTTTCA 481  
 Db GGAACAGAACCCAGGGGTTCTGGCTGTATACACTCTGTTCCTCAAGTTTCCCAAGTTTCA 531  
 QY 482 GCTATGAGTACCCAACTCCGGGAGGCGATATGACAAACCGGATCTGCATCACT- 539  
 Db GCTATGAGTACCCAACTCCGGGAGGCGATATGACAAACCGGATCTGCATCACT- 591  
 QY 540 CGGATACAGCCAGTTT-GAGGAATAGATTTTTCAGCTCTTGGTCTTACCAACCAATCT 598  
 Db CGGATACAGCCAGTTTGGAGGAATAGATTTTTCAGCTCTTGGTCTTACCAACCAATCT 651  
 QY 599 ACAGTCTGCTTCATTCCTCAAGCAGCCTTTTCACTGGAACTTATCTACA 643  
 Db ACAGTCTGCTTCATTCCTCAAGCAGCCTTTTCACTGGAACTTATCTACA 695

RESULT 3  
 AI526873/c  
 LOCUS  
 DEFINITION u147d10.x1 Sugano mouse liver m1ia Mus musculus cDNA clone  
 IMAGE:1923091 3' similar to FR:00115 00115 HYPOTHETICAL HUMAN  
 PROTEIN R31240\_2.; mRNA sequence.  
 AI526873  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus (house mouse)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Marra, M., Hallier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,  
 Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, J.,

person.B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R.,  
Ritter, E., Kohn, S., Shin, T., Jackson, F., Cardenas, M., McCann, R.,  
Waterston, R. and Wilson, R.  
The WashU-NCI Mouse EST Project 1999  
Unpublished (1999)  
Contact: Marra M/WashU-NCI Mouse EST Project 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:979383  
Seq primer: custom primer used  
High quality sequence stop: 490.  
Location/Qualifiers  
1. .510  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL"  
/db\_xref="taxon:10090"  
/clone="IMAGE:1923091"  
/sex="female"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="Sugano mouse liver mlia"  
/notes="Organ: liver; Vector: pME188-FL3; Site: 1: DraIII  
(CACTGTGTG); Site 2: DraIII (CACTGTGTG); 1st strand cDNA  
was primed with an oligo(dT) primer  
[ATGTGGCTCTTTTCTTTTCTTTT]; double-stranded cDNA was  
ligated to a DraIII adaptor [TGTGGCCACTGG], digested  
and cloned into distinct DraIII sites of the pME188-FL3  
vector (5' site CACTGTGTG, 3' site CACTGTGTG). XhoI should  
be used to isolate the cDNA insert. Size selection was  
performed to exclude fragments <1.5kb. Library  
constructed by Dr. Sumio Sugano (University of Tokyo  
Institute of Medical Science). Custom primers for  
sequencing: 5' end primer CTTCGTCTTAAAGCTGGG and 3' end  
primer CGACCTGCAGCTCGAGCACA."  
ORIGIN  
Query Match 41.7%; Score 510; DB 9; Length 510;  
Best Local Similarity 100.0%; Pred. No. 1.5e-148; Indels 0; Gaps 0;  
Matches 510; Conservative 0; Mismatches 0;  
715 CAGGGCTAAACCTTCGTCATTTTGCATAATCAAGTTTATATCTGATGACATTTTACA 774  
710 CAGGGCTAAACCTTCGTCATTTTGCATAATCAAGTTTATATCTGATGACATTTTACA 451  
775 GGATGGATAGCTCAAAAGTTGAAGACACATTTGTAGCACAACCTGGCGAAGAA 834  
770 GGATGGATAGCTCAAAAGTTGAAGACACATTTGTAGCACAACCTGGCGAAGAA 391  
835 CAAGAGCTTCCTTCAAACTGTTCCTGCTTACCATGTCTACACATCAAGTCCATTTGGG 894  
390 CAAGAGCTTCCTTCAAACTGTTCCTGCTTACCATGTCTACACATCAAGTCCATTTGGG 331  
895 GTAACTTCCAAAGTCTTACTTCAGTTTCGCGAAGACCATTTCCAAATGGTGTTCATA 954  
330 GTAACTTCCAAAGTCTTACTTCAGTTTCGCGAAGACCATTTCCAAATGGTGTTCATA 271  
955 AAGGGCTCCGAAATCCGCTGCACTGTGAGACCTTAATCAAGCTACACCAAGCC 1014  
270 AAGGGCTCCGAAATCCGCTGCACTGTGAGACCTTAATCAAGCTACACCAAGCC 211  
1015 TTAAGAGGTGAGGATTCATCTGTACAAAGATCACTACATTTACCAGGCAATTTCAAAA 1074  
210 TTAAGAGGTGAGGATTCATCTGTACAAAGATCACTACATTTACCAGGCAATTTCAAAA 151  
1075 TTATATCTCCGTTATGGGTTCTGTAAAGTAACTCGTGGAAGGGCACACCTCTGTCTTT 1134  
150 TTATATCTCCGTTATGGGTTCTGTAAAGTAACTCGTGGAAGGGCACACCTCTGTCTTT 91

1135 GAAACACTGGCACTGGACATCTCCCTTGGATCTGTCTCCATAATCAAGGCTTCT 1194  
90 GAAACACTGGCACTGGACATCTCCCTTGGATCTGTCTCCATAATCAAGGCTTCT 31  
1195 GAGTGAGCACAACGCTAGCGTCCAAATAAAG 1224  
30 GAGTGAGCACAACGCTAGCGTCCAAATAAAG 1  
RESULT 4  
BQ681950 893 bp mRNA linear EST 15-JUL-2002  
LOCUS AGENCOURT\_8183335 NIH\_MGC\_112 Homo sapiens cDNA clone IMAGE:6262400  
DEFINITION 5', mRNA Sequence.  
ACCESSION BQ681950  
VERSION BQ681950.1 GI:21794629  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 893)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: DCTD/DTF  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LLC2424 row: k column: 09  
High quality sequence stop: 647.  
Location/Qualifiers  
1. .893  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6262400"  
/tissue\_type="melanotic melanoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 112"  
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Library constructed by Ling Hong in the  
Laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH MGC Library."  
ORIGIN  
Query Match 39.6%; Score 485.2; DB 13; Length 893;  
Best Local Similarity 74.7%; Pred. No. 1.2e-140; Indels 6; Gaps 1;  
Matches 623; Conservative 0; Mismatches 205;  
356 ACGATGTGTCCCTGGATCTGTGAATTCAGCAGACAGTATGGACATGCCAAAGTCTGC 415  
1 ATGATGGAGTCCCTTAAACCTGTGAATTCAGCAGACAGTATGGACATGCCAAAGTCTGC 60  
416 TGGTATGACACAGACCGCAGGGTTCGGCTGATACACTCTGTTCCTCCAGTTTCCCCCAG 475  
61 TCGTGTGGAACAGAGTTCAGGGTTCGGCTGATTCATTCATCCCTCAGTTTCTCCAA 120  
476 TTC-----ATGGCTATGAGTACCCAACTCGGGAGCGGATATGACAAACCGGATCT 529  
121 TTCGGGAAGAGGCTATGATTTATCCACCCACAGGAGACGAAATGGAACAAAGTGGCATCT 180  
530 GCATCACTTCGGGATACAGCCAGTTTGAGGAAATAGATTTTCAGCTCTTGCTTTACAAC 589

```

181 GCATAACTTTCAAGTACAAACAGTATGAGGCAATAGATTCTCAGCTTGTGTCGAACC 240
590 CAAACATCTACAGCTGCTTCAATCCAGACACCTTCTCAGTGAACATTAATCATGCCCC 649
241 CCAAGCTCTATAGCTGCTCCATCCAGCCACCTTTCCACAGGAGCTCAATCATGCCCC 300
650 GGATGTGTCCAACTCCAGTTCCTTAAAGATCCCTGTCGCGTACCTCGCTGAACCTGCACT 709
301 AGCTGTGACACAGGCGCCAGCTCATCAGAGATTCTGCGCAGGCTCCTCACCACACTTCAGT 360
710 CAGCCACAGGTCWAACCTGCTCCATTTGCAAAATCAAGTTTATTAATGATGACATCT 769
361 CGGCCCCAGGACAAAATTCCTCCATTTGCAAGTCGGATTCTTTCTTGACGACATCT 420
770 TTACAGGATGATAGCTCAAAAGTTGAAGACACATTTGTAGACAAACCTGCGCAAAA 829
421 TTGACGCTGGATGGCTCAACGCTGAAGACACATTTGTAACAGAAACCTGCGACGAA 480
830 AGAAACAGAGCTTCCTTCAAACTGTTCCCTGCTTACCATGCTCAACATCAAGTCCA 889
481 AAGACACAGAGCTTCTTCAAACTGCTCCCTTCTTACCATGCTCAATATAAAGCAA 540
890 TTGGGGTAATCTCCAGTCTTACTTCAAGTCTGCGCAAGACCATTCCAAATGGTGTGTT 949
541 TTAAATATTATCAGACACTCTTATTTCAAGTCTTATCAAGATCAAGCAAGTGGTGTATT 600
950 CCATAAAGGCTCCGCAAAATCGCTGACCTGCAATGGAGACCTTAATCAAGCTTACAC 1009
601 CCCAAAGGCGACAAAATCGCTGACATGATTTGGAGACCTTANATCGGATGCCACAC 660
1010 AAGCCTTAAGAGTGGAGGATTCATCTGTACAAAGATCACTACATTTACCGAGCATTC 1069
661 AAGCCTTCAGAGTGGAGGATTCATTTGTACCAAGATTTGCAAAATTTACCAAGATTN 720
1070 ATAAATTAATCTCGTTATGGTGTCTGTAACTGCTGGAAGGCGCACACCTCTG 1129
721 CAGGATTATGATTAATGATGAAAGCTGTAACTGCTGGAAGGCGCACAGGTACTA 780
1130 TCTTTGAAACACCTGCACTGCAATCTCGCTTGGATCTGTTTCCATAATA 1183
781 TCATTGAAACCTTGACAAATGGGTCTCTTCCATTACACCTTCTTTATATTTTA 834

```

```

RESULT 5
BX526247
LOCUS
DEFINITION
BX526247 Sugano mouse liver mlia Mus musculus cDNA clone
EST.
BX526247.1 GI:32297078
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 546)
Heil, O., Ebert, L., Neubert, P., Peters, M., Radelof, U., Schneider, D.
and Korn, B.
Mouse Unigeneset - RZPD2
Unpublished (2003)
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGE998M204718;
RZPD; IMAGE998M204718;
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
http://www.rzpd.de/cloncards/cgi-
bin/showlib.pl.cgi/response?libNo=981
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111

```

# www.rzpd.de This clone is available royalty-free from RZPD; contact RZPD (clone@rzpd.de) for further information. Seq primer: sugf, Primer sequence: CTTCTGCTCTAAAGCTGCG. Location/Qualifiers 1. 546 /organism="Mus musculus" /mol\_type="mRNA" /strain="C57BL" /db\_xref="taxon:10090" /clone="IMAGE998M204718 ; IMAGE:1923643" /sex="female" /dev\_stage="adult" /lab\_host="DH10B" /clone\_lib="Sugano mouse liver mlia" /note="Organ: liver; Vector: pME18-FL3; Site 1: DraIII (CACTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCTACTGG], digested and cloned into distinct DraIII sites of the pME18-FL3 vector (5' site CACTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTAAAGCTGCG and 3' end primer CGACTGCGAGCTCGAGACA."

## FEATURES

### source

## ORIGIN

```

Query Match 37.5%; Score 459.4; DB 13; Length 546;
Best Local Similarity 99.8%; Pred. No. 1.3e-132;
Matches 460; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCAGTCCCTCATGGAATGAAGCCACAGATAGAAAATGACAGCAAGCCCTCTAAGAA 61
DB 86 CACAGTCCCTCATGGAATGAAGCCACAGATAGAAAATGACAGCAAGCCCTCTAAGAA 145
QY 62 CAGTTCTTTCTTTGCTCTTTCTTTGCTCTCTGCGGTCTGCGGACACAGAAATCTCAT 121
DB 146 CAGTTCTTTCTTTGCTCTTTCTTTGCTCTCTGCGGTCTGCGGACACAGAAATCTCAT 205
QY 122 CGAGAAATGAATATGCTGAAGCTGGAAGCTGTTTATCTTTTATAGTTACCAAGAA 181
DB 206 CGAGAAATGAATATGCTGAAGCTGGAAGCTGTTTATCTTTTATAGTTACCAAGAA 265
QY 182 CTAGCAAGCGCAAGTGAAGAGCGGGCTGCACTACTCTGCTGACTCCCAAGACAAA 241
DB 266 CTAGCAAGCGCAAGTGAAGAGCGGGCTGCACTACTCTGCTGACTCCCAAGACAAA 325
QY 242 CTTGGAACAAGAGCTCTTACCTGATTAACAGACAGAGTGCTCTGCGGAGGACCTTAC 301
DB 326 CTTGGAACAAGAGCTCTTACCTGATTAACAGACAGAGTGCTCTGCGGAGGACCTTAC 385
QY 302 AGCATCTGTATGACACATATTCACGATGACAGACCTTCTATATATACAGATG 361
DB 386 AGCATCTGTATGACACATATTCACGATGACAGACCTTCTATATATACAGATG 445
QY 362 GTGTCCTTGGATCTCTGAAATTAACAGACAGATGATGACATGCCAAGGCTGCTGGTAT 421
DB 446 GTGTCCTTGGATCTCTGAAATTAACAGACAGATGATGACATGCCAAGGCTGCTGGTAT 505
QY 422 GGAACAGAACCGAGGGTCTGCGTGTATACACTCTGTTTCCC 462
DB 506 GGAACAGAACCGAGGGTCTGCGTGTATACACTCTGTTTCCC 546

```

## RESULT 6

### LOCUS

### DEFINITION

AI875954 432 bp mRNA linear EST 21-JUL-1999  
 u153b10.y1 Sugano mouse liver mlia Mus musculus cDNA clone  
 IMAGE:1923643 5' similar to TR:090759 O90759 DEOXYRIBONUCLEASE II  
 ; mRNA sequence.

QY	242	CCTCGAACAAGAGCCTCTACCTGATTTACAGCACACAGAGTGTCTCTGGGAGACCTTAC	301			
Db	260	CTTGGAAACAAGAGCCTCTACCTGATTTACAGCACACAGAGTGTCTCTGGGAGACCTTAC	319			
QY	302	AGCATCTGTATGACACACATATTTCCAGATGACAGAGCTATCTAATATACACAGATG	361			
Db	320	AGCATCTGTATGACACACATATTTCCAGATGACAGAGCTATCTAATATACACAGATG	379			
QY	362	GTGTCCTCGATCTGTGAATTTACAGCACAGATGATGACATGCCAAAGGCTG	414			
Db	380	GTGTCCTCGATCTGTGAATTTACAGCACAGATGATGACATGCCAAAGGCTG	432			
RESULT 7						
CD365871/c						
LOCUS	CD365871	758 bp	mRNA linear EST 29-MAY-2003			
DEFINITION	UI-H-FT2-bjk-p-12-0-UI.s1 NCI CGAP_F12 Homo sapiens cDNA clone					
ACCESSION	UI-H-FT2-bjk-p-12-0-UI 3', mRNA sequence.					
VERSION	CD365871					
KEYWORDS	CD365871.1 GI:31149961					
SOURCE	EST.					
ORGANISM	Homo sapiens (human)					
REFERENCE	Homo sapiens					
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
JOURNAL	NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .					
COMMENT	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index					
Unpublished (1997)						
Contact: Robert Strausberg, Ph.D.						
Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a>						
Tissue Procurement: Dr. Gary W. Hunninghake, U of I						
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa						
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa						
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa						
Clone Distribution: Distribution information can be found at						
<a href="http://genome.uiowa.edu/distribution/cgap.html">http://genome.uiowa.edu/distribution/cgap.html</a>						
Seq primer: M3 FORWARD						
POLYA=Yes.						
FEATURES						
Location/Qualifiers						
1..758						
/organism="Homo sapiens"						
/mol_type="mRNA"						
/db_xref="taxon:9606"						
/clone="UI-H-FT2-bjk-p-12-0-UI"						
/tissue_type="Aveolar Macrophage"						
/dev_stage="Adult"						
/lab_host="DH10B (Life Technologies)"						
/clone_lib="NCI_CGAP_F12"						
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP_F12 is a subcloned cDNA library constructed from a pool of 81 RNA samples from Alveolar Macrophages challenged with different treatments. The library was subcloned according to Bonaldo, Lennon and Soares, Genom Research, 6:791-806, 1996. The tissue was provided by Dr. Gary W. Hunninghake of the University of Iowa.						
TAG TISSUE=Human Lung Avascular Macrophage						
TAG LIB=UI-H-FT2						
TAG_SEQ=GGCCATGCGC						
ORIGIN						
Query Match 32.7%; Score 400; DB 14; Length 758;						
Best Local Similarity 75.0%; Pred. No. 6.8e-114;						
Matches 513; Conservative 0; Mismatches 170; Indels 1; Gaps 1						
QY	500	CGGGGAGCGGATGAGCAAAACCGCATCTGCATCTTCGATACCTTTCCGATACAGCCAGTTTGAGG	559			
Db	757	CAGGAGCAAGCAATGAGCAAA-CTGGATCTGCATCTTCGATACCTTTCCGATACAGCCAGTTTGAGG	699			
QY	560	AAATAGATTTTCAGCTTTGGTCTTCAACCAACCAACATCTACAGCTGCTTCATTCGAGCA	619			

Db 698 CAATAGATTCTAGCTCTTGGTCTGCAACCCCAACGCTCTATATAGCTGCTCCATCCAGCCA 639  
 QY 620 CCTTTCACTGGAACATTATATACATGCCCCGGATGTCGCAACTCCAGTCTCTTAAGA 679  
 Db 638 CCTTTACACAGGAGCTCATTCATGCCCCAGCTGTGCACAGGCGCAGCTCATCAGAGA 579  
 QY 680 TCCTGTCCGGTACCTGCGTGAATCTGCACTCAGCCAGGGCTCTAAATCTTCTCCATTTTG 739  
 Db 578 TTCTGTGCGAGCTCCTCACCACACTTCAGTCGCGCCAGGAGCAAAAATCTCTCCATTTTG 519  
 QY 740 CAAAAATCAAGTTTATATACATGACATCTTTACAGATGATAGCTCAAAAGTTGAAGA 799  
 Db 518 CAAAGTCGGATCTTTCTTGTGATGATCTTTTCAGCCTGGATGGCTCAACGGCTGAAGA 459  
 QY 800 CACATTTGTTAGCACAAACCTGCGACAAAAGAAACAAAGAGCTTCTTCAAACTGTTCCC 859  
 Db 458 CACACTTTGTTAAGAGAACTTGCAGCGGAAAAGACAGAGCTTCTTCAAACTGTTCCC 399  
 QY 860 TGGCTTACCATGCTCAACATCAATGCTCCATTTGGGTAATCTTCAAGTCTTACTTCAGTT 919  
 Db 398 TTCTTACCATGCTCAATAATAAAGCAATTAATATCAAGACACTCTTATTTTCAGTT 339  
 QY 920 CTCGCCAAGACCATTCCTCAATGCTGTTTCCATTAAGGGCTCCGCAATTCGCTGACCT 979  
 Db 338 CTATCAAGATCATGCAAGTGTGTATTTCCAAAAGGGCCACAAAATTCGCTGACAT 279  
 QY 980 GCATTGGAGACCTAAATCAAGCCTTACACCAAGCCTTAAGAGGTGAGGATTCATCTGA 1039  
 Db 278 GTATTGGAGACCTAAATCGAGTCCACACCAAGCCTTCAGAGTGGAGGATTCATTTGA 219  
 QY 1040 CAAAGATCACTACATTTACAGGCAATTCATAAATATATATCTTCGTTATGGGTTCTGA 1099  
 Db 218 CCCAGAAATGGCAAAATTTACCAAGCAATTTCAATGATTAGTATTATATGAAGCTGA 159  
 QY 1100 ACTAACTCGGTGAAGGCCACACCTCTGCTTGAAGAACTGCACTGCAATGCAATCTC 1159  
 Db 158 AGTAACTTGGTGAAGGACACAGGTACTATCATTTGAAAACCTTGCAATGGGCTCTTCT 99  
 QY 1160 GCCTTGGATCTGTTCTCCATATA 1183  
 Db 98 CCATTACACCTCTTTATATTITA 75

RESULT 8  
 LOCUS AI048641 543 bp mRNA linear EST 08-JUL-1998  
 DEFINITION ud64c11.y1 Sugano mouse liver mlia Mus musculus cDNA clone  
 IMAGE:1450676.5, similar to TR:000115 000115 HYPOTHETICAL HUMAN  
 PROTEIN R31240\_2. i, mRNA sequence.

ACCESSION AI048641  
 VERSION AI048641.1 GI:3296928  
 KEYWORDS Mus musculus (house mouse)  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (Bases 1 to 543)  
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,  
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and  
 Waterston, R.  
 TITLE The WashU-HMI Mouse EST Project  
 JOURNAL Unpublished (1996)  
 COMMENT Contact: Marra M/Mouse EST Project  
 WashU-HMI Mouse EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@wustl.edu  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:923992  
 Seq primer: custom primer used  
 High quality sequence stop: 518.  
 Location/Qualifiers  
 1. 543  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:1450676"  
 /sex="female"  
 /dev stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="Sugano mouse liver mlia"  
 /note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII  
 (CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA  
 was primed with an oligo(dT) primer  
 [ATGTCCTTTTCTTTTCTTTTCTTTT]; double-stranded cDNA was  
 ligated to a DraIII adaptor [GTGTGCTTACTGG], digested  
 and cloned into distinct DraIII sites of the pME18S-FL3  
 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should  
 be used to isolate the cDNA insert. Size selection was  
 performed to exclude fragments <1.5kb. Library  
 constructed by Dr. Sumio Sugano (University of Tokyo  
 Institute of Medical Science). Custom primers for  
 sequencing: 5' end primer CTCTGCTCTTAAAGCTGCG and 3' end  
 primer CGACCTGAGCTCGAGACA."

FEATURES  
 source  
 1. 543  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:1450676"  
 /sex="female"  
 /dev stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="Sugano mouse liver mlia"  
 /note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII  
 (CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA  
 was primed with an oligo(dT) primer  
 [ATGTCCTTTTCTTTTCTTTTCTTTT]; double-stranded cDNA was  
 ligated to a DraIII adaptor [GTGTGCTTACTGG], digested  
 and cloned into distinct DraIII sites of the pME18S-FL3  
 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should  
 be used to isolate the cDNA insert. Size selection was  
 performed to exclude fragments <1.5kb. Library  
 constructed by Dr. Sumio Sugano (University of Tokyo  
 Institute of Medical Science). Custom primers for  
 sequencing: 5' end primer CTCTGCTCTTAAAGCTGCG and 3' end  
 primer CGACCTGAGCTCGAGACA."

ORIGIN  
 Query Match 31.2%; Score 382.4; DB 9; Length 543;  
 Best Local Similarity 99.7%; Pred. No. 2.1e-108;  
 Matches 383; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 CCCAGTCCCCTGCATGGAATGAAGGCCACAGATAGAAAATGACAGCAAGCCTCTAAGAA 61  
 Db 160 CACAGTCCCCTGCATGGAATGAAGGCCACAGATAGAAAATGACAGCAAGCCTCTAAGAA 219  
 QY 62 CAGTCTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 121  
 Db 220 CAGTCTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 279  
 QY 122 CAGAAATGAATATGCTGAGCTGGAAGCTGGAAGCTGGAAGCTGGAAGCTGGAAGCT 181  
 Db 280 GCAGAAATGAATATGCTGAGCTGGAAGCTGGAAGCTGGAAGCTGGAAGCTGGAAGCT 339  
 QY 182 CTAGCAAGGCAAGTGAAGAGGCGGGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 241  
 Db 340 CTAGCAAGGCAAGTGAAGAGGCGGGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 399  
 QY 242 CCTGGAACAAGAGCTCTACCTGATTAACAGACAGAGGAGTCTCTGGGGAGGACCTTAC 301  
 Db 400 CCTGGAACAAGAGCTCTACCTGATTAACAGACAGAGGAGTCTCTGGGGAGGACCTTAC 459  
 QY 302 AGCATCTGTATGACACACATAATTCAGATGACAGACCTCTATCTATATATACAGATG 361  
 Db 460 AGCATCTGTATGACACACATAATTCAGATGACAGACCTCTATCTATATATACAGATG 519  
 QY 362 GTGTCCCTTGGATCTGTGAATTACA 385  
 Db 520 GTGTCCCTTGGATCTGTGAATTACA 543

RESULT 9  
 CD365107/c  
 LOCUS CD365107  
 DEFINITION UI-H-PT2-bjn-1-05-0-UI.s1 NCI CGAP\_Ft2 Homo sapiens cDNA clone  
 UI-H-PT2-bjn-1-05-0-UI.3', mRNA sequence.  
 ACCESSION CD365107  
 VERSION CD365107.1 GI:31149197  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CD365107 729 bp mRNA linear EST 29-MAY-2003  
 UI-H-PT2-bjn-1-05-0-UI.s1 NCI CGAP\_Ft2 Homo sapiens cDNA clone  
 UI-H-PT2-bjn-1-05-0-UI.3', mRNA sequence.  
 ACCESSION CD365107  
 VERSION CD365107.1 GI:31149197  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

RESULT 9  
 CD365107/c  
 LOCUS CD365107  
 DEFINITION UI-H-PT2-bjn-1-05-0-UI.s1 NCI CGAP\_Ft2 Homo sapiens cDNA clone  
 UI-H-PT2-bjn-1-05-0-UI.3', mRNA sequence.  
 ACCESSION CD365107  
 VERSION CD365107.1 GI:31149197  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

307 AAAAAATCGCTGCAGCATGTATTGGAGACCTTAATCGGAGTCCACACCAAGCCTTCAGAGT 248  
1024 GGAGGATTCATCTGTACAAAGAAATCACTACATTTACCAGGCAATTTCAATAATATATCTC 1083  
247 GGAGGATTCATTTGTACCCAGAAATTTGGCAATTTACCAAGCAATTTCAAGATAGTATTA 188  
1084 CGTTATGGGTTCTGTAAAGTAAACTCGTGCAAGGCGACACCCCTCTCTCTCTGAAACACT 1143  
187 TACTATGAAGCTGTAAAGTAAACTTGGTGAAGGACACAGGTACTATCATTTGAAACCTT 128  
1144 GGCACTGGAAACATCTCGCTTGAATCTGTCTCCATAATA 1183  
127 GCAATGGGTTCTTCTCCATTACACCTTCTTATATTTA 88

RESULT 10  
BG761484 923 bp mRNA linear EST 15-MAY-2001  
LOCUS 602718694F1 NTH\_MGC\_49 Homo sapiens cDNA clone IMAGE:485894 5',  
DEFINITION mRNA sequence.  
ACCESSION BG761484 GI:114072137  
VERSION BG761484  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 923)  
NIH-MGC http://mgi.nci.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgabs-remail.nih.gov  
Tissue Procurement: ATCC/DCTD/DTF  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LNCM1713 row: k column: 23  
High quality sequence stop: 845.

FEATURES  
Location/Qualifiers  
1..923  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:485894"  
/tissue\_type="melanotic melanoma, high MDR (cell line)"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 49"  
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; CDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGACGAG(G). Size-selected >500bp for average insert size  
1.8kb. Library constructed by Ling Hong in the laboratory  
of Gerald M. Rubin (University of California, Berkeley)  
using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
II RT (Life Technologies). Note: this is a NIH\_MGC  
Library."

ORIGIN  
Query Match 29.9%; Score 366.4; DB 12; Length 923;  
Best Local Similarity 73.3%; Pred. No. 2.7e-103;  
Matches 597; Conservative 0; Mismatches 191; Indels 26; Gaps 9;  
358 GATGGTTCCTGGATCTGTGAATACAGACAGTATGACATGCCAAGGCTCTGCTG 417  
1 GATGGAGTCCCTAAACCTGTGAATACAGAGAGATGACACACCAAGTTTACTG 60  
418 GTATGGACACGACGAGGGTTCCTGGCTGATACACTCTGTTCCCAAGTTTCCCGAGTT 477  
61 CTGTGGAAACAGATTCAGGGTTCCTGGCTGATTCATTCACCTCCCTCAGTTCTCCAATT 120

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 729)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: Dr. Gary W. Humminghake, U of I  
CDNA library preparation: Dr. M. Bento Soares, University of Iowa  
CDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
http://genome.uiowa.edu/distribution/cgap.html  
Seq primer: M13 FORWARD  
POLYA=Yes.

FEATURES  
Location/Qualifiers  
1..729  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-H-FT2-bjn-1-05-0-UI"  
/tissue\_type="Aveolar Macrophage"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NCI CGAP FT2"  
/notes="Organ: Lung; Vector: p7T73-Pac (Pharmacia) with a  
modified polylinker; Site 1: EcoRI; Site 2: Not I;  
NCI CGAP FT2 is a subcloned cDNA library constructed from  
a pool of 81 RNA samples from Alveolar Macrophages  
challenged with different treatments. The library was  
subtracted according to Bonaldo, Lennon and Soares, Genome  
Research, 6:791-806, 1996. The tissue was provided by Dr.  
Gary W. Humminghake of the University of Iowa.  
TAG TISSUE=Human Lung Aveolar Macrophage  
TAG LIB=UI-H-FT2  
TAG\_SEQ=GGCCATGCGG"

ORIGIN  
Query Match 31.0%; Score 379.2; DB 14; Length 729;  
Best Local Similarity 74.5%; Pred. No. 2.3e-107;  
Matches 477; Conservative 0; Mismatches 163; Indels 0; Gaps 0;  
544 TACAGCAGTTTGGAGAAATAGATTTTCAGCTCTTGGCTTACAAACCAATCTACAGC 603  
727 TACAACAGTATGAGCAATAGATTTCTCAGCTCTTGGTCTGCAACCCCAAGCTCTATAGC 668  
604 TGCTTCATTCACAGCCTTTCTCACTGGAACTTATACATGCCCCCGGATGTGTGCCAAC 663  
667 TGCTTCATTCACAGCCTTTCTCACTGGAGCTCATTCATGCCCCCGGATGTGTGCCAAC 608  
664 TCCAGTTCCTTAAAGATCCCTGTGCGGTAACCTGCTGACATGCTGACATGCTGACATGCTG 723  
607 GCCAGCTCATCAGAGATTCCTGGCAGGCTCTCACCACACTTCAGTCCGCGCCGAGCAAA 548  
724 AACTTCGCTCAATTTGAAATCAAGTTTATATGATGATGATGATGATGATGATGATGATGAT 783  
547 AAATTCCTCCATTTGAAATCGGATTCCTTCTTGACGACATCTTTCGAGCTGATGATGATG 488  
784 GCTCAAAAGTTGAAGACACATTTGTTAGCAAAACCTGGCAGAAAGAAACAAAGAGCTT 843  
487 GCTCAACGGCTGAAGACACATTTGTTAAGCAAAACCTGGCAGAAAGAAACAAAGAGCTT 428  
844 CTTTCAAACTGTTCCCTGCTTACATGCTTACAAATCATGCTCAATGCTGCTGCTGCTGCTGCT 903  
427 CTTTCAAACTGCTCCCTTCTTACATGCTTACATATATAAAGCAATTAATATATATATATAT 368  
904 AAGTCTTACTTCAGTTCTCGCAACCAATTCCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 963  
367 CACTCTTATTCAGTTCTTATCAAGATCATGCAAGTGTGTTATTTCCAAAGGGCACC 308  
964 GCAATCGCTGGACCTGCTGAGACCTTAATCGAAGCTTACACCAAGCCTTAAGAGGT 1023











919		TCCTGGCAAGACCATTCCTAAAGTGTGTGTTTCCATTAAGGGCTCCCAAAATCGCTGGACC	978
QY			
319		TCTTATCAAGATCATGCCAAGTGTGTATTTTCCAAAGGGCCACCAAAATCGCTGGACA	260
Db			
979		TGCATTGGAGACCTTAATCGAAGCCTTACACCAAGCCTTAAGAGGTGGAGGATTCATCTGT	1038
QY			
259		TGTAITGGAGACCTTAATCGAGTCCACACCAAGCCTTCAGAAGTGGAGGATTCATTTGT	200
Db			
1039		ACAAAGAAATCATCATATTTACCAAGGCATTTCAFAAATTTATCTCCGTTAAGGTTCTGT	1098
QY			
199		ACCCAGAAITGGCAATTTTACCAAGCATTTTCAAGGATTTAGTATTATCATATGAAGCTGT	140
Db			
1099		AGTAAATCGGTGAAGGCCACACCTCTGTCTCTTGAAGAACTGGCACTGGAACATCT	1158
QY			
139		AGTAAATCTGGTGAAGGACACAGGTACTATCATTTGAAACCTTGACAAATGGGTCTTCT	80
Db			
1159		CGCTTTGGATCTGTCTCCATAATATCAAGGCTTCTGAGTGAAGCAACGTAAGCGTCCAA	1218
QY			
79		TCCATTACAC--TTCTTTATATTTTAAAGGCTGTGAATACTTATAACCTGCATATCA	22
Db			
1219		TAAAA 1223	
QY			
21		CAAA 17	
Db			
RESULT 15			
BG149286/c			
LOCUS			
DEFINITION			
BG149286 609 bp mRNA linear EST 05-FEB-2001			
nad25g11.x1 NCI CGAP Lu24 Homo sapiens cDNA clone IMAGE:3366717 3'			
similar to TR:Q9QZK9 Q9QZK9 DEOXYRIBONUCLEASE DLAD. ;, mRNA			
sequence.			
BG149286			
BG149286.1 GI:12661316			
EST.			
Homo sapiens (human)			
Homo sapiens			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Theria; Primates; Catarrhini; Hominoidea; Homo.			
REFERENCE			
1 (bases 1 to 609)			
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.			
AUTHORS			
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),			
Tumor Gene Index			
JOURNAL			
Unpublished (1997)			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgaps-x@mail.nih.gov			
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.			
Emmert-Buck, M.D., Ph.D.			
cDNA Library Preparation: M. Bento Soares, Ph.D.			
DNA Library Arrayed by: Greg Lennon, Ph.D.			
DNA Sequencing by: Washington University Genome Sequencing Center			
Clone distribution: NCI-CGAP clone distribution information can be			
found through the I.M.A.G.E. Consortium/LLNL, send email to:			
info@image.llnl.gov			
Seq primer: -40UP from Gibco			
High quality sequence stop: 416.			
Location/Qualifiers			
1..609			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:3366717"			
/tissue_type="carcinoid"			
/lab_host="DH10B"			
/clone_lib="NCI CGAP Lu24"			
/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a			
modified polylinker; Plasmid DNA from the normalized			
library NCI CGAP Lu5 was prepared, and ss circles were			
made in vitro. Following HAP purification, this DNA was			
used as tracer in a subtractive hybridization reaction.			
The driver was PCR-amplified cDNAs from a pool of 5,000			
clones made from the same library (clonids			
1414920-1417991 and 1520904-1522439). Subtraction by Bento			

Soares and M. Fatima Bonaldo. "			
Query Match		25.0%; Score 305.8; DB 12; Length 609;	
Best Local Similarity		72.3%; Pred. No. 2.7e-84;	
Matches 397; Conservative 0; Mismatches 152; Indels 0; Gaps 0;			
<hr/>			
QY	635	TTATCTACATGCCCGGATGTGTGCCAACTCCAGATTCCCTTAAAGATCCCTGTCGGTACC	694
Db	608	TCATTTCAATGCTCCAGCTGTGCACAGGACGAGCTCACTGAGATCTTGGCAGAGCTCA	549
QY	695	TCGCTGAAATGCACTCAGCCCGGGTCTAAACTTCGTCATTTTGGCAAAATCAAGTTTTT	754
Db	548	TCACCAACAATTGATCGGCCCGGACAAAAATTCCTCAATTTTGGCAAAATCGGATTTCT	489
QY	755	ATACTGATGACATCTTTACAGGATGATAGCTCAAAAGTTGAAAGACACATTTCTTTAGCAC	814
Db	488	TTCTTGATGACATCTTTTGAGGCTTGGATGATCAAGCGCTGAAGACACACTTTGTAACAG	429
QY	815	AAACCTGGCAGAAAAAGAAACAGAGCTTCCCTTCAAACTGTTCCTGCTTACCATGTCT	874
Db	428	AAACCTGGCAGCGAAAAAGACAGAGCTTCCCTTCAAACTGTTCCTTCCCTTACCATGTCT	369
QY	875	ACAACATCAAGTCCATTTGGGGTAACTTCCAAGTCTTACTTCAAGTCTTCGCGCAAGACATT	934
Db	368	ACAATATAAAGCAATTAAATTTATCAGACACTCTTATTTTCACTTCTTATCAAGATCATG	309
QY	935	CGAATGCTGTGTTTCCATTAAGGGCTCCGCAATCGCTGACCTGCATTTGGAGACCTAA	994
Db	308	CCAAGTGGTGTATTTCCCAAAAGGGCACCACCAAAATCGCTGACATGTTTGGAGACCTAA	249
QY	995	ATCGAAGCCTTACACCAAGCCTTAAGAGGTGGAGGATTCATCTGTACAAAGAATCACTACA	1054
Db	248	ATCGAGTCCACACCAAGCCTTCAGAAGTGGAGGATTCATTTTGAACCCAGATTGGCAAA	189
QY	1055	TTTACCAAGCATTTCATAAATTTATCTCCGTTATGCGTTCTGTAGTAAACTCGGTGAA	1114
Db	188	TTTACCAAGCATTTCAGAAGTGTAGTATTATATACTATGAAGCTGTAGTAAACTTGGTGA	129
QY	1115	AGGCCACACCTCTGCTTCTTGAAGAAACACCTGCACTGCAACATCTCGCTTTGGATCTGTC	1174
Db	128	AGGACACAGGTACTATCATTTGAAAAACCTTGACAAATGGGTCTTCTTCCATTACACCTTCTT	69
QY	1175	TCCATAATA 1183	
Db	68	TATATTTTA 60	

Search completed: October 14, 2004, 05:24:09  
Job time : 2992.85 secs

Result No.	Score	Query		DB	ID	Description
		Match	%			
1	1224	100.0	1224	9	US-09-949-434-1	Sequence 1
2	1224	100.0	1224	13	US-10-240-709-1	Sequence 1
3	1224	100.0	1224	17	US-10-790-589-1	Sequence 1
4	1213.8	99.7	1652	17	US-10-670-863-2	Sequence 1
5	672.8	55.0	1268	9	US-09-949-434-3	Sequence 3
6	672.8	55.0	1268	13	US-10-240-709-3	Sequence 3
7	672.8	55.0	1268	17	US-10-790-589-3	Sequence 3
8	641.8	52.4	1086	17	US-10-670-863-4	Sequence 4
9	193.2	15.8	480	10	US-09-918-995-7149	Sequence 4
10	65	5.3	65	10	US-09-908-975-28714	Sequence 4
11	60.8	5.0	1575	16	US-10-408-167A-1	Sequence 1
12	53.8	4.4	65	10	US-09-908-975-3418	Sequence 1
13	48.8	4.0	60	10	US-09-908-975-13222	Sequence 1
14	41	3.3	379	9	US-09-960-352-7234	Sequence 7

	Query Match	Best Local Similarity	100.0.%; Score 1224;	DB 9;	Length 1224;
	Matches 1224;	Conservative	0; Mismatches	0; Indels	0; Gaps
QY	1	TCCAGTCCCCGTCATGGGAATGAAGGCCACACATAGAGAAAATGACAGAAAGCCTCTTAAGA	60		
Db	1	TCCAGTCCCCGTCATGGGAATGAAGGCCACACATAGAGAAAATGACAGAAAGCCTCTTAAGA	60		
QY	61	ACAGTCTCTTTCTTTTGCTCTCTTTTGGCCCTCTCTGGGGTCCCTGGGGACACACAGAAATCTCA	120		
Db	61	ACAGTCTCTTTCTTTTGCTCTCTTTTGGCCCTCTCTGGGGTCCCTGGGGACACACAGAAATCTCA	120		
QY	121	TGCAGAAATGAATATGGTGAAGCTGTGACCTGTGGTTTATCTTTTATAGATTACCCAAAAGG	180		
Db	121	TGCAGAAATGAATATGGTGAAGCTGTGACCTGTGGTTTATCTTTTATAGATTACCCAAAAGG	180		

181 ACTAGCAGGCAAGTGAAGAGCGGGCTGACGTACCTGTACCTGACCTCCACAGACAA 240  
181 ACTAGCAGGCAAGTGAAGAGCGGGCTGACGTACCTGTACCTGACCTCCACAGACAA 240  
241 ACCTGGAACAAGCCTCTACCTGATTAACAGCACCAGAGTGTCTCTGGGAGGACCTTA 300  
241 ACCTGGAACAAGCCTCTACCTGATTAACAGCACCAGAGTGTCTCTGGGAGGACCTTA 300  
301 CAGCATCTGTATGACACACATAATCCACGAATGACACAGCCTATCTAATATACAAAGAT 360  
301 CAGCATCTGTATGACACACATAATCCACGAATGACACAGCCTATCTAATATACAAAGAT 360  
361 GGTGTCCTGATCTGTGATTAACAGCAGATGAGCATGCCCAAGAGTGTCTGCTGTA 420  
361 GGTGTCCTGATCTGTGATTAACAGCAGATGAGCATGCCCAAGAGTGTCTGCTGTA 420  
421 TGGACACAGAACCGAGGGTCTGGCTGATACACTCTGTCCCAAGTTTCCCCAGTTTCA 480  
421 TGGACACAGAACCGAGGGTCTGGCTGATACACTCTGTCCCAAGTTTCCCCAGTTTCA 480  
481 GGCTATGAGTACCACCTCGGGAGCGGATGAGCAAAACCGGCTCTGCTACACTTTTC 540  
481 GGCTATGAGTACCACCTCGGGAGCGGATGAGCAAAACCGGCTCTGCTACACTTTTC 540  
541 GATACAGCAGTTTGAAGAAATAGATTTTTCAGCTCTTGTGCTTACCAACCAACATCTAC 600  
541 GATACAGCAGTTTGAAGAAATAGATTTTTCAGCTCTTGTGCTTACCAACCAACATCTAC 600  
601 AGCTGCTTCAATCCAAAGCACCCTTTCAGTGGAACTTATCTCATGCCCGGATGTGCC 660  
601 AGCTGCTTCAATCCAAAGCACCCTTTCAGTGGAACTTATCTCATGCCCGGATGTGCC 660  
721 CTAAACTTCGTCATTTTGAAATCAAGTTTATTAATGATGATGATGATGATGATGATG 780  
721 CTAAACTTCGTCATTTTGAAATCAAGTTTATTAATGATGATGATGATGATGATGATG 780  
781 ATAGCTCAAAAGTTGAAGACACATTTGTTAGCACAAACTGGCAGAAAGAAACAAGAG 840  
781 ATAGCTCAAAAGTTGAAGACACATTTGTTAGCACAAACTGGCAGAAAGAAACAAGAG 840  
841 CTTCTCTCAAACTGTTCCCTGCTTACAGTCTTACAACTCAAGTCAAGTGGGTTAACT 900  
841 CTTCTCTCAAACTGTTCCCTGCTTACAGTCTTACAACTCAAGTCAAGTGGGTTAACT 900  
901 TCCAGTCTTACTTCAAGTCTTCCGCAAGACCATTTCCAAATGGTGTGTTCCATTAAGGC 960  
901 TCCAGTCTTACTTCAAGTCTTCCGCAAGACCATTTCCAAATGGTGTGTTCCATTAAGGC 960  
961 TCCGCAAACTCGTGGACCTGATTTGGAGACCTAAATCGAAGCTTACCAAGCCTTAAGA 1020  
961 TCCGCAAACTCGTGGACCTGATTTGGAGACCTAAATCGAAGCTTACCAAGCCTTAAGA 1020  
1021 GGTGGAGATTCATCTGTACAAAGATCACTACATTTACAGGATTTTCAATAATTATAT 1080  
1021 GGTGGAGATTCATCTGTACAAAGATCACTACATTTACAGGATTTTCAATAATTATAT 1080  
1081 CTCCGTTATGGGTTCTGTAAAGTAACTCGGTGAAGAGGCAACCTCTGTCTTGAAGAAC 1140  
1081 CTCCGTTATGGGTTCTGTAAAGTAACTCGGTGAAGAGGCAACCTCTGTCTTGAAGAAC 1140  
1141 ACTGGCACTGGAACATCTCGCTTGGATCTGTCTTCCATATATCAAGGCTTCTGAGTGA 1200  
1141 ACTGGCACTGGAACATCTCGCTTGGATCTGTCTTCCATATATCAAGGCTTCTGAGTGA 1200  
1201 GCACACGTAGGCTCAATAAAG 1224  
1201 GCACACGTAGGCTCAATAAAG 1224

RESULT 2  
US-10-240-709-1  
; Sequence 1, Application US/10240709  
; Publication NO. US20030212023A1  
; GENERAL INFORMATION:  
; APPLICANT: Eastman, Alan R.  
; APPLICANT: Krieser, Ronald J.  
; APPLICANT: Trustees of Dartmouth College  
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS  
; FILE REFERENCE: DC-0154  
; CURRENT APPLICATION NUMBER: US/10/240,709  
; CURRENT FILING DATE: 2002-10-01  
; PRIOR APPLICATION NUMBER: 09/541,840  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: 09/574,942  
; PRIOR FILING DATE: 2000-05-19  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1224  
; TYPE: DNA  
; ORGANISM: Mus sp.  
US-10-240-709-1  
Query Match 100.0%; Score 1224; DB 13; Length 1224;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCCAGTCCCTGTCATGGAATGAAGCCACAGATAGAAATGACAGCAAGCCCTCTAAGA 60  
DB 1 TCCAGTCCCTGTCATGGAATGAAGCCACAGATAGAAATGACAGCAAGCCCTCTAAGA 60  
QY 61 ACAGTTCTTTCTTTTGTCTTTCTTTTGTCTCTCTGGGCTCTCTGGGACACCAAGAAATCTCA 120  
DB 61 ACAGTTCTTTCTTTTGTCTTTCTTTTGTCTCTCTGGGCTCTCTGGGACACCAAGAAATCTCA 120  
QY 121 TGCAGAAATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180  
DB 121 TGCAGAAATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180  
QY 181 ACTAGCAGGCAAGTGAAGAGCGGGCTGACGTACCTGTACCTGACCTCCACAGACAA 240  
DB 181 ACTAGCAGGCAAGTGAAGAGCGGGCTGACGTACCTGTACCTGACCTCCACAGACAA 240  
QY 241 ACCTGGAACAAGCCTCTACCTGATTAACAGCACCAGAGTGTCTCTGGGAGGACCTTA 300  
DB 241 ACCTGGAACAAGCCTCTACCTGATTAACAGCACCAGAGTGTCTCTGGGAGGACCTTA 300  
QY 301 CAGCATCTGTATGACACACATAATCCACGAATGACACAGCCTATCTAATATACAAAGAT 360  
DB 301 CAGCATCTGTATGACACACATAATCCACGAATGACACAGCCTATCTAATATACAAAGAT 360  
QY 361 GGTGTCCTGATCTGTGATTAACAGCAGATGAGCATGCCCAAGAGTGTCTGCTGTA 420  
DB 361 GGTGTCCTGATCTGTGATTAACAGCAGATGAGCATGCCCAAGAGTGTCTGCTGTA 420  
QY 421 TGGAAACAGAACCGAGGGTCTGGCTGATACACTCTGTCCCAAGTTTCCCCAGTTTCA 480  
DB 421 TGGAAACAGAACCGAGGGTCTGGCTGATACACTCTGTCCCAAGTTTCCCCAGTTTCA 480  
QY 481 GGCTATGAGTACCACCTCGGGAGCGGATGAGCAAAACCGGCTCTGCTACACTTTTC 540  
DB 481 GGCTATGAGTACCACCTCGGGAGCGGATGAGCAAAACCGGCTCTGCTACACTTTTC 540  
QY 541 GGATACAGCAGTTTGAAGAAATAGATTTTTCAGCTCTTGTGCTTACCAACCAACATCTAC 600  
DB 541 GGATACAGCAGTTTGAAGAAATAGATTTTTCAGCTCTTGTGCTTACCAACCAACATCTAC 600  
QY 601 AGCTGCTTCAATCCAAAGCACCCTTTCAGTGGAACTTATCTCATGCCCGGATGTGCC 660  
DB 601 AGCTGCTTCAATCCAAAGCACCCTTTCAGTGGAACTTATCTCATGCCCGGATGTGCC 660  
QY 661 AACTCCAGTTTCTTAAAGATCCCTGTCGGTGAACCTCGTGAACCTGACCTCAGCCAGGGT 720

```

661 AACCTCAGTTCCTTAAAGATCCCTGTCGGTACCTGCTGAACTGCACTCAGCCAGGGT 720
721 CTAATACTTCCTCATTTTGCATAATCAAGTTTATATATATATATATATATATATATAT 780
721 CTAATACTTCCTCATTTTGCATAATCAAGTTTATATATATATATATATATATATATAT 780
781 ATAGCTCAAAAGTTGAAGACACATTTTGTAGCAAAAACCTTGGCAGAAAAGAAACAAGAG 840
781 ATAGCTCAAAAGTTGAAGACACATTTTGTAGCAAAAACCTTGGCAGAAAAGAAACAAGAG 840
841 CTTCCCTTCAAACTGTTCCCTGCTTACCATCTGCTAATCAATCAAGTCCATTTGGGTAACT 900
841 CTTCCCTTCAAACTGTTCCCTGCTTACCATCTGCTAATCAATCAAGTCCATTTGGGTAACT 900
901 TCCAACTCTTACTTCACTGTTCTGCCCAAGACCATTCCTCAATGTTGTTTCCATTAAGGGC 960
901 TCCAACTCTTACTTCACTGTTCTGCCCAAGACCATTCCTCAATGTTGTTTCCATTAAGGGC 960
961 TCCGCAAAATCGCTGACCTGCTTACCTTGAAGACCTTAAATCGAAGCTTACACCAAGCCTTAAAG 1020
961 TCCGCAAAATCGCTGACCTGCTTACCTTGAAGACCTTAAATCGAAGCTTACACCAAGCCTTAAAG 1020
1021 GGTGAGGAGTTCATCTGTACAAAGATCACTACATTTTACAGGCAATTCATTAATATATAT 1080
1021 GGTGAGGAGTTCATCTGTACAAAGATCACTACATTTTACAGGCAATTCATTAATATATAT 1080
1081 CTTCCGTTATGGTTCCTGTAAGTAACTCGGTGAAAGGCCACACCTCTGCTCTTGAAGAAC 1140
1081 CTTCCGTTATGGTTCCTGTAAGTAACTCGGTGAAAGGCCACACCTCTGCTCTTGAAGAAC 1140
1141 ACTGGCACTGGAACATCTCGCTTGGATCTGTTCTTCCATAATCAAGGCTTCTGAGTGA 1200
1141 ACTGGCACTGGAACATCTCGCTTGGATCTGTTCTTCCATAATCAAGGCTTCTGAGTGA 1200
1201 GCACAACGTAGCGTCCATAAAG 1224
1201 GCACAACGTAGCGTCCATAAAG 1224

```

```

RESULT 3
US-10-790-589-1
; Sequence 1, Application US/10790589
; Publication No. US20040142376A1
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Trustees of Dartmouth College
; TITLE OF INVENTION: DEXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCE: DC-0154
; CURRENT APPLICATION NUMBER: US/10790,589
; PRIOR FILING DATE: 2004-03-01
; PRIOR FILING DATE: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; PRIOR FILING DATE: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1224
; TYPE: DNA
; ORGANISM: Mus sp.
US-10-790-589-1

```

```

Query Match 100.0%; Score 1224; DB 17; Length 1224;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

61 ACAGTTCTTTCTTTGCTCTTTTGGCCCTCTCTGGGGTCTCTGGGGACACAGAAATCTCA 120
121 TGCAGAAATGAATATATATATATATATATATATATATATATATATATATATATATATAT 180
121 TGCAGAAATGAATATATATATATATATATATATATATATATATATATATATATATATAT 180
181 ACTAGCAAGCAAGTCAAGAGCGGGGCTGCACTCTCTGATCTGCTGACTCCACAAGACAA 240
181 ACTAGCAAGCAAGTCAAGAGCGGGGCTGCACTCTCTGATCTGCTGACTCCACAAGACAA 240
241 ACCTGGAACAAGAGCTCTTACCTGATTAACAGAGCGGGGCTGCACTCTCTGATCTGCTG 300
241 ACCTGGAACAAGAGCTCTTACCTGATTAACAGAGCGGGGCTGCACTCTCTGATCTGCTG 300
301 CAGCATCTGTATGACACACATATATCCAGAAATGACACAGCCTATCTAATATATCAACGAT 360
301 CAGCATCTGTATGACACACATATATCCAGAAATGACACAGCCTATCTAATATATCAACGAT 360
361 GGTGTCCTTGGATCTGTGAAATACAGCAGACAGTATGACATGCAAGGCTCTGCTGTTA 420
361 GGTGTCCTTGGATCTGTGAAATACAGCAGACAGTATGACATGCAAGGCTCTGCTGTTA 420
421 TGGAAACAGAACGAGGCTCTGCTGATACATCTGTTCCCAAGTCTTCCCAAGTCTTCA 480
421 TGGAAACAGAACGAGGCTCTGCTGATACATCTGTTCCCAAGTCTTCCCAAGTCTTCA 480
481 GGTATGAGTACCCAACTCGGGGAGGCGATATGACAAACCGGCTATCTGATCACTTTC 540
481 GGTATGAGTACCCAACTCGGGGAGGCGATATGACAAACCGGCTATCTGATCACTTTC 540
541 GGATACAGCAGTTTGAAGAAATAGATTTTCACTGCTCTTGGTCTTACCAACCAATCTAC 600
541 GGATACAGCAGTTTGAAGAAATAGATTTTCACTGCTCTTGGTCTTACCAACCAATCTAC 600
601 AGCTGCTTCTTCCAAAGCACCTTTCACCTGGAACCTTCTGCTGAACTGCACTCAGCCAG 660
601 AGCTGCTTCTTCCAAAGCACCTTTCACCTGGAACCTTCTGCTGAACTGCACTCAGCCAG 660
661 AACTCCAGTTCCTTAAAGATCCCTGTCGGTACCTGCTGAACTGCACTCAGCCAGGCT 720
661 AACTCCAGTTCCTTAAAGATCCCTGTCGGTACCTGCTGAACTGCACTCAGCCAGGCT 720
721 CTAAGCTTCGTCATTTTGAAGAAATCAAGTTTATATATATATATATATATATATATAT 780
721 CTAAGCTTCGTCATTTTGAAGAAATCAAGTTTATATATATATATATATATATATATAT 780
781 ATAGCTCAAAAGTTGAAGACACATTTTGTAGCAAAAACCTTGGCAGAAAAGAAACAAGAG 840
781 ATAGCTCAAAAGTTGAAGACACATTTTGTAGCAAAAACCTTGGCAGAAAAGAAACAAGAG 840
841 CTTCCCTTCAAACTGTTCCCTGCTTACCATCTGCTAATCAATCAAGTCCATTTGGGTAACT 900
841 CTTCCCTTCAAACTGTTCCCTGCTTACCATCTGCTAATCAATCAAGTCCATTTGGGTAACT 900
901 TCCAACTCTTACTTCACTGTTCTGCCCAAGACCATTCCTCAATGTTGTTTCCATTAAGGGC 960
901 TCCAACTCTTACTTCACTGTTCTGCCCAAGACCATTCCTCAATGTTGTTTCCATTAAGGGC 960
961 TCCGCAAAATCGCTGACCTGCTTACCTTGAAGACCTTAAATCGAAGCTTACACCAAGCCTTAAAG 1020
961 TCCGCAAAATCGCTGACCTGCTTACCTTGAAGACCTTAAATCGAAGCTTACACCAAGCCTTAAAG 1020
1021 GGTGAGGAGTTCATCTGTACAAAGATCACTACATTTTACAGGCAATTCATTAATATATAT 1080
1021 GGTGAGGAGTTCATCTGTACAAAGATCACTACATTTTACAGGCAATTCATTAATATATAT 1080
1081 CTTCCGTTATGGTTCCTGTAAGTAACTCGGTGAAAGGCCACACCTCTGCTCTTGAAGAAC 1140
1081 CTTCCGTTATGGTTCCTGTAAGTAACTCGGTGAAAGGCCACACCTCTGCTCTTGAAGAAC 1140
1141 ACTGGCACTGGAACATCTCGCTTGGATCTGTTCTTCCATAATCAAGGCTTCTGAGTGA 1200

```

Db 1141 ACTGGCACTGGAAACATCTCGCCTGGATCTGTTCTCCATAATATCAAGGCTTCTGAGTGA 1200

QY 1201 GCACAACTAGCGTCCCAATAAAG 1224

Db 1201 GCACAACTAGCGTCCCAATAAAG 1224

RESULT 4

US-10-670-863-2

; Sequence 2, Application US/10670863

; Publication No. US20040157239A1

; GENERAL INFORMATION:

; APPLICANT: Tanuma, Sei-ichi

; APPLICANT: Shiohawa, Daisuke

; TITLE OF INVENTION: Novel Deoxyribonuclease, Gene Encoding Thereof and Use Thereof

; FILE REFERENCE: 224589

; CURRENT APPLICATION NUMBER: US/10/670,863

; CURRENT FILING DATE: 2003-09-25

; PRIOR FILING DATE: JP 11-230870

; PRIOR FILING DATE: 1999-08-17

; PRIOR APPLICATION NUMBER: US 09/807,784

; PRIOR FILING DATE: 2001-04-17

; NUMBER OF SEQ ID NOS: 15

; SEQ ID NO 2

; LENGTH: 1652

; TYPE: DNA

; ORGANISM: Mus Musculus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (213)..(1274)

US-10-670-863-2

Query Match 99.7%; Score 1219.8; DB 17; Length 1652;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1221; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCAGTCCCTGATGATGAGGCGCACAGATAGAAATGACAGCAAGCCCTTAAGAA 61

Db 175 CACAGTCCCTGATGATGAGGCGCACAGATAGAAATGACAGCAAGCCCTTAAGAA 234

QY 62 CAGTCTCTTTCTGCTCTCTTTGCCCCCTCTCTGGGCTCTGGGACACCAAGAAATCTCAT 121

Db 235 CAGTCTCTTTCTGCTCTCTTTGCCCCCTCTCTGGGCTCTGGGACACCAAGAAATCTCAT 294

QY 122 GCAGAAATGAATAGTGAAGCTGTGACCTGTTTATCTTTTATAGTTACCAAGGA 181

Db 295 GCAGAAATGAATAGTGAAGCTGTGACCTGTTTATCTTTTATAGTTACCAAGGA 354

QY 182 CTAGCAAGCAAGTGAAGGCGGGCTGCAGTACCTGTACCTGACCTCCACAGACAAA 241

Db 355 CTAGCAAGCAAGTGAAGGCGGGCTGCAGTACCTGTACCTGACCTCCACAGACAAA 414

QY 242 CCTGGAACAAGAGCCTCTACTGATTAACAGCACAGAGTGTCTCTGGGAGGACCTTAC 301

Db 415 CCTGGAACAAGAGCCTCTACTGATTAACAGCACAGAGTGTCTCTGGGAGGACCTTAC 474

QY 302 AGCATCTGTATGACACATATATTCACAGATGACAGACCTATCTAATATACACCATG 361

Db 475 AGCATCTGTATGACACATATATTCACAGATGACAGACCTATCTAATATACACCATG 534

QY 362 GTGTCCTGTGATCTGTGAATTTACAGCAGACAGTATGACATGCCAAAGGCTCTGCTGAT 421

Db 535 GTGTCCTGTGATCTGTGAATTTACAGCAGACAGTATGACATGCCAAAGGCTCTGCTGAT 594

QY 422 GGAACAGAACCGAGGGTCTGTGGTGTATACACTCTGTTTCCAAAGTTCCTCCAGTTCATG 481

Db 595 GGAACAGAACCGAGGGTCTGTGGTGTATACACTCTGTTTCCAAAGTTCCTCCAGTTCATG 654

QY 482 GCTATGATGATCCCAACCTCGGGAGGCGATATGACAAACCGCATCTGCATCTCTTCG 541

Db 655 GCTATGATGATCCCAACCTCGGGAGGCGATATGACAAACCGCATCTGCATCTCTTCG 714

QY 542 GATACAGCCAGTTTGGAGGAATAGATTTTTCAGCTCTTGGTCTTACACCAACATCTACA 601

Db 715 GATACAGCCAGTTTTCAGGAAATAGATTTTCAGCTCTTGGTCTTACAAACCAATCTACA 774

QY 602 GCTGCTTCAATTCACAGCAACCTTTCACCTGGAACCTTATCTACATGCCCGGATGTGTGCA 661

Db 775 GCTGCTTCAATTCACAGCAACCTTTCACCTGGAACCTTATCTACATGCCCGGATGTGTGCA 834

QY 662 ACTCCAGTTTCCCTTAAAGATCCCTGTCGGGTACCTCGCTGAACCTGCACTCAGCCCGGTC 721

Db 835 ACTCCAGTTTCCCTTAAAGATCCCTGTCGGGTACCTCGCTGAACCTGCACTCAGCCCGGTC 894

QY 722 TAACTTCTGTCATTTTGCAGAAATCAAGTTTTCATGATGATCATCTTTTACAGATGCA 781

Db 895 TAACTTCTGTCATTTTGCAGAAATCAAGTTTTCATGATGATCATCTTTTACAGATGCA 954

QY 782 TAGCTCAAAGTTGAGACACATTTTGTAGCAACAACTGGCAGAAAAAGAAACCAAGAGC 841

Db 955 TAGCTCAAAGTTGAGACACATTTTGTAGCAACAACTGGCAGAAAAAGAAACCAAGAGC 1014

QY 842 TTCCCTTCAAACCTGTTCCCTTACCATGTCTACAAATCAAGTCCATTTGGGGTAACTT 901

Db 1015 TTCCCTTCAAACCTGTTCCCTTACCATGTCTACAAATCAAGTCCATTTGGGGTAACTT 1074

QY 902 CCAAGTCTTACTTCACTTCTGCGCAAGACCATTTCCAAATGCTGTTTCCATAAAGGGCT 961

Db 1075 CCAAGTCTTACTTCACTTCTGCGCAAGACCATTTCCAAATGCTGTTTCCATAAAGGGCT 1134

QY 962 CGGCAATCGCTGGACCTCATTTGGAGACCTAAATCGAAGCCCTTACCAAGCCTTAAAG 1021

Db 1135 CGGCAATCGCTGGACCTCATTTGGAGACCTAAATCGAAGCCCTTACCAAGCCTTAAAG 1194

QY 1022 GTGGAGGATTCATCTGTACAAAGATTCATCTACATTTTACAGGATTTTCAATTAATATC 1081

Db 1195 GTGGAGGATTCATCTGTACAAAGATTCATCTACATTTTACAGGATTTTCAATTAATATC 1254

QY 1082 TCGTATGAGTGTCTGTAAAGTAACTCGGTGAAGGCCACACCTCTGCTCTTGAAGAA 1141

Db 1255 TCGTATGAGTGTCTGTAAAGTAACTCGGTGAAGGCCACACCTCTGCTCTTGAAGAA 1314

QY 1142 CTGGCACTGGAACATCTCGCTTGGATCTGTTCTCCATATATCAAGGCTTCTGAGTGA 1201

Db 1315 CTGGCACTGGAACATCTCGCTTGGATCTGTTCTCCATATTTTCAAGGCTTCTGAGTGA 1374

QY 1202 CACAACTAGCGTCCCAATAAAG 1224

Db 1375 CACAACTAGCGTCCCAATAAAG 1397

RESULT 5

US-09-949-434-3

; Sequence 3, Application US/09949434

; Patent No. US2002028495A1

; GENERAL INFORMATION:

; APPLICANT: Eastman, Alan R.

; APPLICANT: Krieser, Ronald J.

; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS

; FILE REFERENCE: DC-0137

; CURRENT APPLICATION NUMBER: US/09/949,434

; CURRENT FILING DATE: 2001-09-07

; PRIOR FILING DATE: 09/574,942

; PRIOR FILING DATE: 2000-05-19

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3

; LENGTH: 1268

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-949-434-3

Query Match 55.0%; Score 672.8; DB 9; Length 1268;

Best Local Similarity 75.0%; Pred. No. 5.2e-208;

Matches 871; Conservative 0; Mismatches 282; Indels 9; Gaps 2;

QY 31 AGATAGAAATGACAGCAAGCCCTCTAAGAACAGTCTTTCTTTGCTCTCTCTTTGCCCTC 90  
Db 37 AAACAGAAATGATGGCAAGACTGCTAAGAACATCCCTTTGCTTTGCTCTCTCTTTGCCCTC 96  
QY 91 TCTGGGCTCTGGGACACACAGAAATCTCATGAGAAATGAATATGATGAGCTGTGGAC 150  
Db 97 TTTGGGCTCTGGGCGCGCAACAAATTTCTGAGAAATGAAGAAAGGAAAGCTGTGGAC 156  
QY 151 TGGTTTATCTTTTATAGTTTACCAAGGACTAGCAAGGCAAGTGAAGAGGGGGGCTG 210  
Db 157 TGGTTTATCTTTTATAGTTTACCTAAGGCAAAACAGGAAAGTGAGAGACTGGGTTA 216  
QY 211 CAGTACTGTACTGAGCTCCACAGCAAACTGGAAACAGAGCCCTCTACCTGATTAAC 270  
Db 217 GAGTACTGTACTGAGCTCTCAACTAGAACTGGAGGAAGAGTGAGCAACTAATGAAT 276  
QY 271 AGCACCAGAGAGTCTCTGGGAGGACCTTACAGCATCTGTATGACACACATATTCACG 330  
Db 277 GACACCAAGAGTGTGGGAGGAGCAATTAACACAGCTATATGAAGCATATGCTCTAAG 336  
QY 331 AATG---ACACAGCCCTATCTAATATACACGATGGTGTCCCTGGATCTGTGAATTAACG 387  
Db 337 AGTAACACACAGCCCTCTAATATACATGATGGAGTCCCTAAACCTGTGAATTAAGT 396  
QY 388 AGCAGTATGGACATGCCAAAGTCTGTGGTATGGAACAGAACGAGGGGTTCTGGCTG 447  
Db 397 AGAAGTATGGACACACCAAGAGTTTACTGTGTGGAACAGAGTTCAAGGGTTCTGGCTG 456  
QY 448 ATACACTCTGTCCCAAGTTTCCCCAGTTC-----ATGGCTATGAGTACCCACCTCG 501  
Db 457 ATTCAATCCATCCCTCAGTTTCTCCATTTCCGGAAGAGGCTATGATTCACCCACA 516  
QY 502 GGGAGGCGATATGGACAAACCGGCATCTGCATCACTTTCCGATACAGCCAGTTTCAGGAA 561  
Db 517 GGGAGCGAATATGGACAAAGTGGCTCTGCAATACTTTCAAGTACACAGCATGAGGCA 576  
QY 562 ATAGATTTTACGCTCTTGGTCTTACAAACACATCTACAGCTGTCTCAATCCAAAGCACC 621  
Db 577 ATAGATTTTACGCTCTTGGTCTGCAACCCCAACGCTATAGCTGTCCATCCACGCCACC 636  
QY 622 TTTCACTGGAACTTATCTACATGCCCCGGATGTGCGCACTTCTGCTCACTTCTTAAAGATC 681  
Db 637 TTTCAACAGGAGCTCATCTACATGCCCCCAGCTGTGCAACAGGCCAGCTCATCAGAGAT 696  
QY 682 CTTGTCGGTACCTGGCTGAACTGCACTCAGCCCAAGGCTCTAAACTCTGCTCAATTTGCA 741  
Db 697 CTTGCGAGCTCTTCAACCACTTCACTGCGGCCAGGAGCAAAATCTCTCAATTTGCA 756  
QY 742 AAATCAAGTTTATATCTGATGACATCTTTACAGGATGATAGCTCAAAAGTTGAAGACA 801  
Db 757 AAGTCGGATTTCTTTCTTGAGCATCTTTGCGAGCTGATGGCTCAACGGCTGAAGACA 816  
QY 802 CATTTGTTAGCAGAACTCTGACAGAAAGAAACAGAGCTTCTTCAACTGTTCCTG 861  
Db 817 CACTTGTAAACAGAACTCTGACGCGGAAAGAAAGCAAGAGCTTCTTCAAACTGCTCCCT 876  
QY 862 CTTTCAATGFTTCAACATCAAGTCCATTTGGGGTAACTTCCAAAGTCTTACTTCAGTTCT 921  
Db 877 CTTTACCATGCTCAATATAAAGCAATTAATATACAGACACTCTTATTTCAAGTTCT 936  
QY 922 CGCCAAGACCAATTCAAATGTGTGTTTCCATAAAGGGCTCCGCAATTCGCTGACCTGC 981  
Db 937 TATCAAGATCAAGCTGAGTGTGTTATTTCCAAAAGGGCACCAAAATTCGCTGACATGT 996  
QY 982 ATTGGAGACTTAATCGAAGCTTACACCAAGCCCTTAAGAGTGGAGATTCATCTGTACA 1041  
Db 997 ATTGGAGCTTAATTCGAGTTCACCAAGCCCTTCAAGAGTGGAGATTCATTTGTACC 1056  
QY 1042 AAGAATCACTACATTTACCGGCAATTTCAAAATTAATCTCCGTTATGGTCTGTAAAG 1101  
Db 1057 CAGAAATGGCAATTTACCAAGCAATTCAGAGTATGATTAATTAATGAAGCTGTAAAG 1116  
QY 1102 TAAACTCGGTAAAGGCCACACCCCTGTCTCTTTGAAACAACTGGCACTGGAACTCTCGC 1161

Db 1117 TAAACTTGGTGAAGGACACAGGTACTATCATTTGAAACCTTGACATGGTCTTCTTCC 1176  
QY 1162 CTTGGATCTCTTCTCCATAATA 1183  
Db 1177 ATTACACCTCTTTATATTTTA 1198  
RESULT 6  
US-10-240-709-3  
; Sequence 3, Application US/10240709  
; Publication No. US20030212023A1  
; GENERAL INFORMATION:  
; APPLICANT: Eastman, Alan R.  
; APPLICANT: Krieser, Ronald J.  
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS  
; FILE REFERENCE: DC-0154  
; CURRENT APPLICATION NUMBER: US/10/240,709  
; PRIORITY FILING DATE: 2002-10-01  
; PRIOR APPLICATION NUMBER: 09/541,840  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: 09/574,942  
; PRIOR FILING DATE: 2000-05-19  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 1268  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-240-709-3  
Query Match 55.0%; Score 672.8; DB 13; Length 1268;  
Best Local Similarity 75.0%; Pred. No. 5.2e-208;  
Matches 871; Conservative 0; Mismatches 282; Indels 9; Gaps 2;  
QY 31 AGATAGAAATGACAGCAAGCCCTCTAAGAACAGTCTTTCTTTGCTCTCTCTTTGCCCTC 90  
Db 37 AAACAGAAATGATGGCAAGACTGCTAAGAACATCCCTTTGCTTTGCTCTCTTTGCCCTC 96  
QY 91 TCTGGGCTCTGGGACACACAGAAATCTCATGAGAAATGAATATGATGAGCTGTGGAC 150  
Db 97 TTTGGGCTCTGGGCGCGCAACAAATTTCTGAGAAATGAAGAAAGGAAAGCTGTGGAC 156  
QY 151 TGGTTTATCTTTTATAGTTTACCAAGGACTAGCAAGGCAAGTGAAGAGGGGGCTG 210  
Db 157 TGGTTTATCTTTTATAGTTTACCTAAGGCAAAACAGGAAAGTGAGAGACTGGGTTA 216  
QY 211 CAGTACTGTACTGAGCTCCACAGCAAACTGGAAACAGAGCCCTCTACCTGATTAAC 270  
Db 217 GAGTACTGTACTGAGCTCTCAACTAGAACTGGAGGAAGAGTGAGCAACTAATGAAT 276  
QY 271 AGCACCAGAGAGTCTCTGGGAGGACCTTACAGCATCTGTATGACACACATATTCACG 330  
Db 277 GACACCAAGAGTGTGGGAGGAGCAATTAACAGCTATATGAAGCATATGCTCTAAG 336  
QY 331 AATG---ACACAGCCCTATCTAATATACACGATGGTGTCCCTGGATCTGTGAATTAAC 387  
Db 337 AGTAACACACAGCCCTCTAATATACATGATGGAGTCCCTAAACCTGTGAATTAAGT 396  
QY 388 AGCAGTATGGACATGCCAAAGTCTGTGGTATGGAACAGAACGAGGGGTTCTGGCTG 447  
Db 397 AGAAGTATGGACACACCAAGAGTTTACTGTGTGGAACAGAGTTCAAGGGTTCTGGCTG 456  
QY 448 ATACACTCTGTCCCAAGTTTCCCCAGTTC-----ATGGCTATGAGTACCCACCTCG 501  
Db 457 ATTCAATCCATCCCTCAGTTTCTCCATTTCCGGAAGAGGCTATGATTCACCCACA 516  
QY 502 GGGAGGCGATATGGACAAACCGGCATCTGCATCACTTTCCGATACAGCCAGTTTCAGGAA 561  
Db 517 GGGAGCGAATATGGACAAAGTGGCTCTGCAATACTTTCAAGTACACAGCATGAGGCA 576  
QY 562 ATAGATTTTACGCTCTTGGTCTTACAAACACATCTACAGCTGTCTCAATCCAAAGCACC 621  
Db 577 ATAGATTTTACGCTCTTGGTCTGCAACCCCAACGCTATAGCTGTCCATCCACGCCACC 636  
QY 622 TTTCACTGGAACTTATCTACATGCCCCGGATGTGCGCACTTCTGCTCACTTCTTAAAGATC 681  
Db 637 TTTCAACAGGAGCTCATCTACATGCCCCCAGCTGTGCAACAGGCCAGCTCATCAGAGAT 696  
QY 682 CTTGTCGGTACCTGGCTGAACTGCACTCAGCCCAAGGCTCTAAACTCTGCTCAATTTGCA 741  
Db 697 CTTGCGAGCTCTTCAACCACTTCACTGCGGCCAGGAGCAAAATCTCTCAATTTGCA 756  
QY 742 AAATCAAGTTTATATCTGATGACATCTTTACAGGATGATAGCTCAAAAGTTGAAGACA 801  
Db 757 AAGTCGGATTTCTTTCTTGAGCATCTTTGCGAGCTGATGGCTCAACGGCTGAAGACA 816  
QY 802 CATTTGTTAGCAGAACTCTGACAGAAAGAAACAGAGCTTCTTCAACTGTTCCTG 861  
Db 817 CACTTGTAAACAGAACTCTGACGCGGAAAGAAAGCAAGAGCTTCTTCAAACTGCTCCCT 876  
QY 862 CTTTCAATGFTTCAACATCAAGTCCATTTGGGGTAACTTCCAAAGTCTTACTTCAGTTCT 921  
Db 877 CTTTACCATGCTCAATATAAAGCAATTAATATACAGACACTCTTATTTCAAGTTCT 936  
QY 922 CGCCAAGACCAATTCAAATGTGTGTTTCCATAAAGGGCTCCGCAATTCGCTGACCTGC 981  
Db 937 TATCAAGATCAAGCTGAGTGTGTTATTTCCAAAAGGGCACCAAAATTCGCTGACATGT 996  
QY 982 ATTGGAGACTTAATCGAAGCTTACACCAAGCCCTTAAGAGTGGAGATTCATCTGTACA 1041  
Db 997 ATTGGAGCTTAATTCGAGTTCACCAAGCCCTTCAAGAGTGGAGATTCATTTGTACC 1056  
QY 1042 AAGAATCACTACATTTACCGGCAATTTCAAAATTAATCTCCGTTATGGTCTGTAAAG 1101  
Db 1057 CAGAAATGGCAATTTACCAAGCAATTCAGAGTATGATTAATTAATGAAGCTGTAAAG 1116  
QY 1102 TAAACTCGGTAAAGGCCACACCCCTGTCTCTTTGAAACAACTGGCACTGGAACTCTCGC 1161



577	ATAGATTCTCAGCTCTTGCTGTGCAACCGCAACGCTATATAGTGTCTCAATCCAGCCACC	636
622	TTTTCACGTGGAAACCTTATCTACATGCCCCGAGATGTGTGCCAACTTCAGATTTCCCTTAAAGATC	681
637	TTTTCACGAGAGCTCATTTCAATGCCAGAGCTGTGCACACAGGGCCAGCTCATCAGAGATT	696
682	CCTGTCCGGTACCTTCGCTGTAAGTGCATCAGCCCGAGGGTCTTAACTTCGTCCTATTGTGA	741
697	CCTGGCAGGCTCCTCACCACTTCAGTCGGCCGAGGACAAAAATTCCTCCATTGTGA	756
742	AAATCAAGTTTTTATCTATGTATGATCACTTTTACAGGATGGATAGCTCAAAAGTTGAAGACA	801
757	AAGTCGGAATCTTTTTCTTTGACGACATCTTTTCGAGCCTGGATGGCTCAACGGCTGAAGACA	816
802	CATTGTGTAGCACAAACCTGGCGAAAGAAAGAAACAAGAGCTTCCTTCAAACCTGTTCCCTG	861
817	CAC TTGTTTACAGAAACCTGGCAGCGAAAGACAAGAGCTTCCTTCAAACCTGCTCCCTT	876
862	CCTTTCATAGTCTACATCATCAAGTCAATTCATGGGGTAACTTCCAAAGTCTTACTTCAGTCT	921
877	CCTTTACCATGTCTCAATATATAAAGCAATTAATAATATACAGACACTCTTATTTTCAGTTCT	936
922	CGCCAGACCACTTCCAAATGTGTGTGTTTCCATAAAGGGCTCCGCAATCGCTCGAACCTGC	981
937	TATCAAGATACGCCCAAGTGTGTATTTCCAAAGGGCCACCAAAATCGCTGGACATGT	996
982	ATTGGAGACCTTAATTCGAAGCCTACCAACGCCTTAAGAGGTGGAGGATTCATCTGTACA	1041
997	ATTGGAGACCTTAATTCGGAGTCCACACCAAGCCTTCAGAAATGGAGGATTCATTTGTACC	1056
1042	AAGAAATCACTACATTTACAGGCAATTCATAAATATATCTCCGTTATGGTTCTGTGAAG	1101
1057	CAGATTGGCAAAATTCACCAAGCATTTCAAGGATTAGTATTATCTATGAAGAGCTGTGAAG	1116
1102	TAAACTCCGTTGAAGGGCCACACCCCTCTGCTCTTGAAACCACTGGCACTGGAAACATCTCGC	1161
1117	TAAACTTGGTGAAGGACACAGGTACATCATTTGAAACCTTGGACAAATGGGTCTTCTTCC	1176
1162	CTTGGATCTGTTCTCCATAATA	1183
1177	ATTACACCTTCTTTATATTTTA	1198

## RESULT 7

```

US-10-790-589-3
; Sequence 3, Application US/10790589
; Publication No. US20040142376A1
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Krieser, Ronald J.
; APPLICANT: Trustees of Dartmouth College
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS
; FILE REFERENCE: DC-0154
; CURRENT APPLICATION NUMBER: US/10/790.589
; CURRENT FILING DATE: 2004-03-01
; PRIOR APPLICATION NUMBER: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1268
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-790-589-3

Query Match      55.0%; Score 672.8; DB 17; Length 1268;
Best Local Similarity 75.0%; Pred. No. 5.2e-208;
Matches 871; Conservative 0; Mismatches 282; Indels 9; Gaps 2

          31 AGATGAAATAAGCAGCAAGCGCTTAAGAACAGTCTTTCTTTGCTCTCTTTGGCCCTC 90

```



Db 1117 TAAACTTGGTGAAGGACACAGGTACTATCAATGAAAACTTGACAATGGGTCTTCTTCC 117

QY 1162 CTTGGATCTGTTCTCCATAATA 1183

Db 1177 ATTACACCTCTCTTATATTTTA 1198

RESULT 8

US-10-670-863-4

; Sequence 4, Application US/10670863

; Publication No. US20040157239A1

; GENERAL INFORMATION:

; APPLICANT: Tanuma, Sei-ichi

; APPLICANT: Shokawa, Daisuke

; TITLE OF INVENTION: Novel Deoxyribonuclease, Gene Encoding Thereof and Use Th

; FILE REFERENCE: 224589

; CURRENT APPLICATION NUMBER: US/10/670,863

; CURRENT FILING DATE: 2003-09-25

; PRIOR APPLICATION NUMBER: JP 11-230870

; PRIOR FILING DATE: 1999-08-17

; PRIOR APPLICATION NUMBER: US 09/807,784

; PRIOR FILING DATE: 2001-04-17

; NUMBER OF SEQ ID NOS: 15

; SEQ ID NO 4

; LENGTH: 1086

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(1083)

US-10-670-863-4

Query Match 52.4%; Score 641.8; DB 17; Length 1086;

Best Local Similarity 75.9%; Pred. No. 6.2e-198;

Matches 822; Conservative 0; Mismatches 252; Indels 9; Gaps 2;

QY 31 AGATAGAAATGACGCAAGACCTCTTAAGAACAGTCTCTTTTGTCTCTTTTGTCTCTTTTGGCCCTC 90

Db 4 AAACAGAAATGATGGCAAGACTGCTGAAGAACATCCTTTGCTTTTGTCTTTCTTGGCCTC 63

QY 91 TCTGGGTCTCTGGGACACCAGAAATCTCATGCAAGAAATGATATGTTGAAGCTGTGGAC 150

Db 64 TTTGGGTGCTGGGCGACGCAAAATTTATGCGAAATGAAGAAGGGAAGCTGTGGAC 123

QY 151 TGTGTTATCTTTTATAGTTTACCAGAAAGGACTAGCAAGCAAGTCAAGAGCGGGGGCTG 210

Db 124 TGTGTTTACTTTTATAGTTTACCTAAAGACAAACACAGAAAGTGGAGAGACTGGGTTA 183

QY 211 CAGTACTGTACTGACTGCTCCACAGACAAACCTGGAACAGAGCCTCTACCTGATTAAAC 270

Db 184 GAGTACTGTACTGACTCTACAACCTAGAGCTGAGGAAGTGAAGCACTAATGAAT 243

QY 271 AGCACCAGAGTGTCTCTGGGAGGACCTTACAGCATCTGTATGACACACATATTCACG 330

Db 244 GACACCAAGAGTGTTTTGGGAAGGACATTACACAGCTATATGAAGCATATGCTCTAAG 303

QY 331 ---AATGACACAGCCTTATCTAATATACACAGATGCTGCTCCCTGGATCTGTGAATTTACAGC 387

Db 304 AGTTAACACACAGCCTTCTAATATACATGATGAGTCCCTTAACCTGTGAATTCACG 363

QY 388 AGACAGTATGGACATGCCMAAGTCTGTGGPATGGAAACAGAACGCAAGGGTCTTGCTG 447

Db 364 AGAAGATATGGACACACCAAGGTTTACTGCTGTGGAACAGAGTTTCAAGGGTCTGCTG 423

QY 448 ATACACTCTGTTCCCAAGTTTCCCGAGTTC-----ATGGCTATGATGATCCCAACCTCG 501

Db 424 ATTCACTCAATCCCTCAGTTTCTCCAATTCGGGAAGAGGCTATGATTATCCACCCACA 483

QY 502 GGGAGCGATATGGACAAACCGGCATCTGCATCACTTTTCGGATACAGCCAGTTTGGAA 561

Db 484 GGGAGCGAAATGGACAAAGTGGCATCTGCATTAATTTCAAGTACACACCATATGAGCA 543

QY 562 ATAGATTTTCAGCTCTTGCTGTTTACACCAACATCTACAGCTGCTTCATTTCCAGACCC 621

391 CAGTATGACATGCAAGGCTCTGCTGATGAAACAGACGACGAGGCTTCTGGCTGATA 450  
231 AAGTATGACACACCAAGGTTTACTGCTGTGGAACAGAGTTCAAGGTTCTGGCTGATT 290  
451 CACTCTGTCTCCCAAGTTTCCCCAGTTC-----ATGGCTATGAGTACCCAACTCGGGG 504  
291 CATTCATCCCTCAGTTTCTCTCAATTCGGGAAGAGCTATGATTATCCACCCACAGG 350  
505 AGCGGATATGGAACAAACCGGCACTGTCATCTTCGGATACAGCCAGTTTGGGAATA 564  
351 AGACGAATGGAACAAAGTGGCATCTGCATAAATCTTCAAGTACAACAGATGAGGCAATA 410  
565 GATTTTCAGCTCTTGGTCTTACAAACAACTACAGCTGCTTCAATCCAAAGCACCTTT 624  
411 GATTCAGCTCTTGGTCTGCAACCCCAAGCTATAGCTTCTCCATCCAGGACCTTT 470  
625 CACTGGAA 632  
471 CACCAGGA 478  
RESULT 10  
US-09-908-975-28714  
; Sequence 28714, Application US/09908975  
; Publication No. US20030165843A1  
; GENERAL INFORMATION:  
; APPLICANT: SHOSHAN, Avi  
; APPLICANT: WASSERMAN, Alon  
; APPLICANT: MINTZ, Eli  
; APPLICANT: MINTZ, Liat  
; APPLICANT: FAIGLER, Simchon  
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE  
; FILE REFERENCE: 36688-0005  
; CURRENT APPLICATION NUMBER: US/09/908,975  
; CURRENT FILING DATE: 2001-07-20  
; PRIOR APPLICATION NUMBER: US 60/287,724  
; PRIOR FILING DATE: 2001-05-02  
; PRIOR APPLICATION NUMBER: US 60/221,607  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 32337  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 28714  
; LENGTH: 65  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-908-975-28714  
Query Match 5.3%; Score 65; DB 10; Length 65;  
Best Local Similarity 100.0%; Pred. No. 1.7e-10;  
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1118 CCACACCTCTGTCTCTTGAACAACTGGCACTGGAACATCTCGCTTGGATCTCTTCC 1177  
Db 1 CCACACCTCTGTCTCTTGAACAACTGGCACTGGAACATCTCGCTTGGATCTCTTCC 60  
QY 1178 ATAAAT 1182  
Db 61 ATAAAT 65  
RESULT 11  
US-10-408-167A-1  
; Sequence 1, Application US/10408167A  
; Publication No. US20030219428A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Baron, Will F.  
; TITLE OF INVENTION: Human DNase II  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way

CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/408,167A  
FILING DATE: 04-Apr-2003  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/861034  
FILING DATE: 18-May-2001  
APPLICATION NUMBER: 08/639294  
FILING DATE: 25-Apr-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Evans, David W.  
REGISTRATION NUMBER: NONE  
REFERENCE/DOCKET NUMBER: P1024D1C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-1739  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1575 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: DNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-408-167A-1  
Query Match 5.0%; Score 60.8; DB 16; Length 1575;  
Best Local Similarity 44.8%; Pred. No. 3.2e-08;  
Matches 395; Conservative 0; Mismatches 467; Indels 24; Gaps 3;  
QY 200 AGCGGGGCTGCAGTACTCTGACTGACCTCCAAAGAGTCTGCTGGTATGGAACAGAGCGCTCT 259  
Db 222 AGAGAGGGCTGCAGTACTGACGAGAGTCTGACGAGAGTCTCGGAGGCTGGCGGAGCGGAGG 281  
QY 260 ACCTGATTAACAGACACAGAGAGTCTTGGGAGGAGCTTACAGCATCTGTATGACACAC 319  
Db 282 CACTCATCAACAGCCCGAGGGGCGCTGGGCGGAGCCCTGCAGCCGCTGTACCGGAGCA 341  
QY 320 ATAAATCCAGATGACACAGCCCTATCTAATATACACAGATGGTGTCCCTGGATCTGTGA 379  
Db 342 ACACAGCCAGCTGCGCTTCTCTGCTCTACATGACCAACCGCCTCAACCCAGCAGGCTC 401  
QY 380 ATTACAGCAGACAGTATGGACATGCGCAAGAGTCTGCTGGTATGGAACAGAGCGGCT 439  
Db 402 AGNACTCTTCATCGTGGGACACAGAGAGTGTCTCTGCTCTTGGACACGATGGGGCT 461  
QY 440 TCTGGCTGATACATCTGTTCCTCCAAAGTTTCCCCAGTCTTCCGCTATGAGTACCCAGCT 499  
Db 462 TCTGGCTGTCCACAGTGTACTAACTTCCCTCCACCGGCGCTCTCTGTGCTGATACAGCT 521  
QY 500 CG-----GGGAGGCGATATGGAACAAACCGGCATCTGCATCACTTTCGGATACAGCC 550  
Db 522 GGCCTCATAGCGCTGTACCTACGGGAGAGCCCTGCTCTGTGTGCTCTTTCCCTTGGCTC 581  
QY 551 AGTTTGGAGGAATAGATTTTTCAGCTCTTGGTCTTGTACCAACAAACATCTACAGCTGTCTCA 610  
Db 582 AGTTCTCGAAGATGGCAAGCAGCTGACCTACACCTACCCCTGGGTCTTATAACTACCCAGC 641  
QY 611 TTCCAGCACCTTTTCACTGGGAACCTTATCTACATGCCCCCGAGTGTGTGCCAACTCCAGTT 670  
Db 642 TGGAGGGGATCTTTGCCCAGGAATTTCCCGAGCTTGGAGAATGTGGTCAAGGGCCACACAG 701  
QY 671 CCTTAAAGATCCCTGTCCGGTACCTCGCTGACTGCACTACAGCCAGGCTTAAACTCG 730

RESULT 13  
US-908-975-13222  
; Sequence 13222, Application US/09908975  
; Publication No. US20030165843A1  
; GENERAL INFORMATION:  
; APPLICANT: SHOSHAN, Avi

Query Match	3.3%;	Score 41;	DB 9;	Length 379;
Best Local Similarity	55.1%;	Pred. NO. 0.037;		
Matches 102;	Conservative 0;	Mismatches 80;	Indels 3;	Gaps 1;
QY	130	GAATATGTTGAAGCTGTGACCTGGTTTACTCTTTTATAAGTTACCCAAAAGGACTGACGAG	189	
Db	118	GACTCGGGGACCCGTGTGACTGGTTGTCGATTACAAAGCTGCGGCCCCACACAGGGTCG	177	
QY	190	GCAAGTG---AAGAGCGGGGCTGCAGTACCTGTACTCTGGACTCCCAAGACACAAACCTGG	246	
Db	178	GGAGATGCGACGACAGACGGCTCTGCGGTACAAGTACTTTTGACGAACACTCATAGAAGCTGG	237	
QY	247	AACAAAGACCTCTACTCTGATTAAACAGCACAGAGTGTCTCTGGGAGGACCTTACAGCAT	306	
Db	238	AGCGACGCGTGGGGGTCACTCAATAGCACCAACCGGTGCCGTGGGCCGCACTGCTGCCG	297	
QY	307	CTGTA	311	
Db	298	CTGTA	302	

Search completed: October 14, 2004, 08:18:15  
Job time : 597.827 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1224	100.0	1224	4	US-09-574-942-1		Sequence 1, Appli
2	1219.8	99.7	1258	4	US-09-807-784B-2		Sequence 2, Appli
3	672.8	55.0	1262	4	US-09-574-942-3		Sequence 3, Appli
4	641.8	52.4	1086	4	US-09-807-784B-4		Sequence 4, Appli
5	60.8	5.0	1575	3	US-08-639-294-1		Sequence 1, Appli
6	60.8	5.0	1575	4	US-09-861-034B-1		Sequence 1, Appli
7	60.8	5.0	1915	3	US-09-147-915-1		Sequence 1, Appli
8	55.4	4.5	927	3	US-09-147-915-2		Sequence 2, Appli
9	36.2	3.0	1230025	4	US-09-138-452A-1		Sequence 1, Appli
10	36	2.9	7218	1	US-08-232-463-14		Sequence 14, Appli
11	35.4	2.9	1005	4	US-09-547-435-7		Sequence 7, Appli
12	35.4	2.9	1383	4	US-09-547-435-11		Sequence 11, Appli
13	35.4	2.9	1470	4	US-09-547-435-11		Sequence 3, Appli
14	35.4	2.9	1848	4	US-09-547-435-9		Sequence 9, Appli
15	35.4	2.9	1938	4	US-09-547-435-25		Sequence 25, Appli
16	35.4	2.9	2236	4	US-09-547-435-5		Sequence 5, Appli
17	35.4	2.9	2316	4	US-09-547-435-27		Sequence 27, Appli
18	35.4	2.9	2604	4	US-09-547-435-23		Sequence 23, Appli
19	35.4	2.9	2701	4	US-09-547-435-1		Sequence 1, Appli
20	35.4	2.9	3384	4	US-09-547-435-29		Sequence 29, Appli
21	35	2.9	1230025	4	US-09-198-452A-1		Sequence 1, Appli
22	34.8	2.8	262	4	US-09-313-294A-2280		Sequence 2280, App
23	33.8	2.8	640681	4	US-09-790-988-1		Sequence 1, Appli
24	33.4	2.7	1828	4	US-09-744-900-2		Sequence 2, Appli
25	33.4	2.7	3742	4	US-09-744-900-1		Sequence 1, Appli
26	33.2	2.7	396	4	US-08-640-173-140		Sequence 140, App
27	33.2	2.7	396	4	US-09-713-550-140		Sequence 140, App

```

; NUMBER OF SEQ ID NOS: 15
; SEQ ID NO 2
; LENGTH: 1652
; TYPE: DNA
; ORGANISM: Mus Musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (213)..(1274)
US-09-807-784B-2

Query Match      99.7%; Score 1219.8; DB 4; Length 1652;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1221; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCCAGTCCCTCCATGGATGAAGCCACAGATAGAAATGACAGCAAAAGCCTCTAAGAA 61
DB 175 CACAGTCCCTCCATGGATGAAGCCACAGATAGAAATGACAGCAAAAGCCTCTAAGAA 234
QY 62 CAGTTCCTTTCTTTTGTCTCTTTTGGGCTCTCTGGGCTCTGGGACACAGAAATCTCAT 121
DB 235 CAGTTCCTTTCTTTTGTCTCTTTTGGGCTCTCTGGGCTCTGGGACACAGAAATCTCAT 294
QY 122 GCAGAAATCAATATGATGTAAGCTGTGGACTGGTTTATCTTTTATAGTTTACCCAAAGGA 181
DB 295 GCAGAAATCAATATGATGTAAGCTGTGGACTGGTTTATCTTTTATAGTTTACCCAAAGGA 354
QY 182 CTAGCAAGCAAGTGAAGAGGGGGCTGCAGTACCTGTACCTGTGGACTCCACAAAGCAA 241
DB 355 CTAGCAAGCAAGTGAAGAGGGGGCTGCAGTACCTGTACCTGTGGACTCCACAAAGCAA 414
QY 242 CTTGGAACCAAGAGCCTCTTACCTGATTAACAGCACAGGAGTGTCTGGGGAGGACCTTAC 301
DB 415 CTTGGAACCAAGAGCCTCTTACCTGATTAACAGCACAGGAGTGTCTGGGGAGGACCTTAC 474
QY 302 AGCATCTGTATGACACACATTAATCCAGAAATGACAGACGCTATCTAATATACAAGATG 361
DB 475 AGCATCTGTATGACACACATTAATCCAGAAATGACAGACGCTATCTAATATACAAGATG 534
QY 362 GTCTCCCTGTATCTGTAATTAACAGACAGATGATGCAATGCCAAGGCTCTCTGTAT 421
DB 535 GTCTCCCTGTATCTGTAATTAACAGACAGATGATGCAATGCCAAGGCTCTCTGTAT 594
QY 422 GGAACAGAACGCGAGGGTTCTGGCTGATACACTCTGTCTCCAAAGTTTCCCCAGTTCAATG 481
DB 595 GGAACAGAACGCGAGGGTTCTGGCTGATACACTCTGTCTCCAAAGTTTCCCCAGTTCAATG 654
QY 482 GCTATGATACCCAACTCTGGGAGGCGATATGGACAAACCGGCATCTGCATCACTTTGG 541
DB 655 GCTATGATACCCAACTCTGGGAGGCGATATGGACAAACCGGCATCTGCATCACTTTGG 714
QY 542 GATACAGCCAGTTTGAGGAAATAGATTTTTCAGCTCTTTGGTCTTACACCAAAATCTTACA 601
DB 715 GATACAGCCAGTTTGAGGAAATAGATTTTTCAGCTCTTTGGTCTTACACCAAAATCTTACA 774
QY 602 GCTGCTTCATTCGAAGCACTTTTCACTGGAAACTTTATCTACATGCCCGGATGTGCCA 661
DB 775 GCTGCTTCATTCGAAGCACTTTTCACTGGAAACTTTATCTACATGCCCGGATGTGCCA 834
QY 662 ACTCCAGTTCTTAAAGATCCCTGTCCGGTACCTCGTGAACCTGCATCCAGCCCGAGGTC 721
DB 835 ACTCCAGTTCTTAAAGATCCCTGTCCGGTACCTCGTGAACCTGCATCCAGCCCGAGGTC 894
QY 722 TAAACTTCGTCCTCAATTTTGGCAAAATAGATTTTTCAGCTCTTTGGTCTTACAGATGGA 781
DB 895 TAAACTTCGTCCTCAATTTTGGCAAAATAGATTTTTCAGCTCTTTGGTCTTACAGATGGA 954
QY 782 TAGCTCAAAAGTTGAAGACACATTTTGTAGCAAACTGGCAGAAAAGAAACAAGGC 841
DB 955 TAGCTCAAAAGTTGAAGACACATTTTGTAGCAAACTGGCAGAAAAGAAACAAGGC 1014
QY 842 TTCCCTCAAAAGTTTCCCTGCTTACCATGCTCTACCAATCAAGTCCATTTGGGGTAACTT 901
DB 1015 TTCCCTCAAAAGTTTCCCTGCTTACCATGCTCTACCAATCAAGTCCATTTGGGGTAACTT 1074
```

```

RESULT 2
US-09-807-784B-2
; Sequence 2, Application US/09807784B
; Patent No. 6653118
; GENERAL INFORMATION:
; APPLICANT: Tanuma, Sei-ichi
; APPLICANT: Shikawa, Daisuke
; TITLE OF INVENTION: No. 6653118el Deoxyribonuclease, Gene Encoding Thereof and Use Th
; FILE REFERENCE: 210792
; CURRENT APPLICATION NUMBER: US/09/807,784B
; CURRENT FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: JP 11-230870
; PRIOR FILING DATE: 1999-08-17
```

337	AGTAAACAACAGAGCCTATCTCTAATATACAAATGATGAGTCCCTAAACCTGTGTAATTACAGT	399
388	AGACAGTATGGACATGCGCAAAAGGTCGTCTGGTATATGAAACAGAACCGACAGGGGTCTCGGCTG	447
397	AGAAAGTATGGACACACCAAGGTTTACTGTCTGTGGAACAGATTCAGAGGTTCTTGGCTG	456
448	ATACACTCTGTCTCCCAAGTTTCCGCCAGTTC-----ATGGCTATGAGTACCCACCTCG	501
457	ATTCAATCCATCCCTCGAGTTTCTCCCAATTCGGAAAGGCTATGATATATCCACCCACA	516
502	GGGAGGCGATATGGACAAAACGGGATCTGCACTCTTTCGGGATACAGCCAGTTTCAGAA	561
517	GGGAGCGAATATGACAAAGTGGCATCTGCAATACCTTCAAGTACAAACAGTATGAGGCA	576
562	ATAGATTTTCAGCTCTTGGTCTTACAAACCAACATCTACAGCTGCTTCATTCCAAGCACC	621
577	ATAGATTTCTCAGCTCTTGGTCTGCAACCCCAACGCTATAGCTGTCTCATCCAGCCACC	636
622	TTTTCACATGGAACTTATCTACATGCCCGGATGTGTGCCAATCCAGTTCCTTAAAGATC	681
637	TTTTCACAGAGGCTCATTCATGCCCGCAGCTGTGCACAGGGCCAGCTCATCGAGATT	696
682	CCTGTCGGGTACCTCGCTGGAATCACTCAGCCCCAGGGTCTAAACTCTCGTCCATTTTGCA	741
697	CCTGGCAGGCTCCTCACACACTTCAGTCGGCCCCAGGGACAAAATTCCTCCATTTTGC	756
742	AAATCAAGTTTATATCTGATGACATCTTTTACAGATGGATAGCTCAAAAGTTGAAGACA	801
757	AAGTCGGATTCTTTTCTTGAAGACATCTTTGACGCTGGATGGCTCAACGGCTGAAGACA	816
802	CATTGTGTTAGCACAAACCTGGCAGAAAAGAAACAGAGCTTCCTTCAACTGTGTCCCTG	861
817	CAC TTGTTTAAAGAACTCGCAGCGAAAAGACAGAGCTTCCTTCAAACTGCTCCCTT	876
862	CCTTACCAATGTACAAACATCAAGTCCAATTGGGGTAACTTCCAAAGTCTTACTTCAGTTCT	921
877	CCTTACCATTGTCTACAAATAAAAGCAATTAATATATCAGACACTCTTATTTCAAGTTCT	936
922	CGCCAGACCAATTCCTCAAAATGGTGTGTTTCATAAAGGGCTCCGCAAAATCGCTGACCTGC	981
937	TATCAAGATCACGCCAAGTGGTGTATTTTCCAAAAGGGCCACAAAATCGCTGACATGT	996
982	ATTGGAGACCTAAATCGAAGCCTACACCAAGCCTTAAGAGGTGAGGATTCATCTGTACACA	1041
997	ATTGGAGACCTAAATCGAGTCCACACCAAGCCTTCAGAAAGTGGAGGATTCATTTGTACC	1056
1042	AAGAATCACTACATTTTACAGGCAATTCATAAATTAATCTCCGTTATGGGTTCTGTAAAG	1101
1057	CAGAAATGGCAAAATTTACCAAGCAATTTCAAGGATTTAGTATTATACTATGAAAGCTGTAAG	1116
1102	TAAACTCGGTGAAGGCCACACCTCTGCTCTGTTGAAAACACTGGCACTGGAAACATCTCGC	1161
1117	TAAACTTGGTGAAGAGACACAGGTACTATCATTTGAAAACTTGACAAATGGGTCTTCTTCC	1176
1162	CTTGGATCTGTTCTCCATAATA	1183
1177	ATTACACCTTCTTATATTTTA	1198

RESOL 7  
US-09-807-784B-4

```

; Sequence 4, Application US/09807784B
; Patent No. 6653118
; GENERAL INFORMATION:
; APPLICANT: Tanuma, Sei-ichi
; APPLICANT: Shiohawa, Daisuke
; TITLE OF INVENTION: No. 6653118el Deoxyribonuclease, Gene Encoding Thereof and Use The
; FILE REFERENCE: 210732
; CURRENT APPLICATION NUMBER: US/09/807,784B
; CURRENT FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: JP 11-230870
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 15

```

```

; SEQ ID NO 4
; LENGTH: 1086
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1083)
US-09-807-784B-4

Query Match      52.4%; Score 641.8; DB 4; Length 1086;
Best Local Similarity 75.9%; Pred. No. 1.9e-210;
Matches 822; Conservative 0; Mismatches 252; Indels 9; Gaps 2;

QY 31 AGATGAAATGACAGCAAGCTCTTAAGACAGTCTTCTTTGCTCTCTTCTTGGCCCTC 90
DB 4 AAACAGAAATGATGGCAAGCTCTTAAGACAGTCTTCTTTGCTCTCTTCTTGGCCCTC 63

QY 91 TCTGGGGTCTGGGGACACAGAAATCTCATGAGAAATGAATATGATGTAAGCTGTGGAC 150
DB 64 TTTGGGGTCTGGGGGAGCAACAATTTCAATGAGAAATGAGAGGAAAGCTGTGGAC 123

QY 151 TGGTTTATCTTTTATAGTTTACCCAAAGGACTAGCAAGCCAAAGTGAAGAGGGGGCTG 210
DB 124 TGGTTTACTTTTATAGTTTACCTTAAAGACAAAACAAGGAAAGTGGAGAGCTGGGTTA 183

QY 211 CAGTACTGTACCTGGATCCCAAGCAAAACCTGGNACAGAGCCTCTACCTGATTAAC 270
DB 184 GAGTACTGTACCTAGACTCTCAACTAGAAAGCTGGAGGAGAGTGAAGCAACTAATGAAT 243

QY 271 AGCAGCAGAGTGTCTCTGGGGAGGACCTTACAGATCTGTATGACACACATAATTCACG 330
DB 244 GACCCAGAGTGTCTTGGGAAGGACATTAACAAGCTATGAGAGTAATGCTCTAAG 303

QY 331 ---AATGACACAGCCTATCTAATATACAAAGATGGTCTCCTGATCTGTAATACAGC 387
DB 304 AGTAACACACAGCCTATCTAATATACAAAGATGGTCTCCTTAAACCTGTGAATACAGC 363

QY 388 AGCAGTATGACATGCCAAAGCTCTGTGTATGAGACAGAGCGAGGTTCTGGCTG 447
DB 364 AGAAAGTATGACACACAAAGGTTTACTGTGTGGAACAGAGTTCAAGGGTTCTGGCTG 423

QY 448 ATACACTGTCTTCCCAAGTTTCCCCAGTTTC-----ATGGCTATGATGATCCCAACCTCG 501
DB 424 ATTCATTCATTCCTCAGTTTCTTCCATTCGGAAGAGGCTATGATATCCACCCACA 483

QY 502 GGGAGGCGATATGACAAACCGGATCTGCATCTTTTGGATACAGCCAGTTTGAGAA 561
DB 484 GGGAGCGAATGACAAAGTGGCATCTGCATACTTTCAAGTACAAACCAAGTATGAGCA 543

QY 562 ATAGATTTTACGCTCTTGGCTTTACACCAACATCTACAGCTCTTCAATTCAGCACC 621
DB 544 ATAGATTTCTAGCTCTTGGCTTGAACCCCAACCTCTATAGCTGTCTCCATCCAGCCACC 603

QY 622 TTTCACTGGAACCTTATCTACATGCCCGGATGTGCAACTCCAGTTCCTTAAAGATC 681
DB 604 TTTCAACAGGAGCTCAITCAGTCCCGCAGCTGTGACAGGCGCAGCTCATCAGATT 663

QY 682 CCTGTCCGGTACCTCGCTGAATCGACTCAGCCCAAGGCTTAAACTTGTCCATTTTGA 741
DB 664 CTGGCAGGCTCTCTCACCACACTTCAGTCCGCCCCAGGAGCAAAAATTCCTCCATTTTGA 723

QY 742 AAATCAAGTTTATATCTGATGACATCTTTACAGATGATAGCTTCAAAAGTGAAGACA 801
DB 724 AAGTCGGATTTCTTTTCTGATGGCACTTTTGAGGCTGGATGGCTCAACGCTGAGACA 783

QY 802 CATTTGTTAGCAGAAACCTGGCAGAAAAGAAAAGAGAGCTTCTTCAAACTGTTCCCTG 861
DB 784 CACTTGTAAACAGAAACCTGGCAGGAAAAGAGAGAGCTTCTTCAAACTGTTCCCTT 843

QY 862 CTTTACCAAGTGTACAAATCAAGTCCATTTGGGGTAACTTCCAGTCTTACTTCAATTCT 921
DB 844 CTTTACCAAGTGTCTACATATAAAGCAATTAATATATACAGATGGTCTCCCTCGATCTGTA 903

922 CGCCAAAGACCAATTCACAAATGTTGTTTCCATAAAGGGTCCGCCAAATCGCTGGACCTGC 981
904 TATCAAGATCATGCCAAAGTGGTATTTCCCAAAAGGGCACCAAAAATCGCTGGACATGT 963
982 ATTGGAGACCTAAATCGAAGCCTACACCAAGCCTTAAGAGTGGAGGATTCATCTGTACA 1041
964 ATTGGAGACCTAAATCGGAGTCCACACCAAGCCTTCAGAAGTGGAGGATTCATTTGTACC 1023
1042 AAGATCATCTACATTTACCAGCATTTCATAAATATATCTCGTTATGGTTCCTCTAAG 1101
1024 CAGAATTTGCCAAATTTACCAAGCATTTCAAGGATTAGTATTATCTATGAAAGCTGTAG 1083
1102 TTA 1104
1084 TTA 1086

RESULT 5
US-08-639-294-1
; Sequence 1, Application US/08639294
; Patent No. 6265195
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Baron, Will F.
; TITLE OF INVENTION: Human DNase II
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/639,294
; FILING DATE: 25-Apr-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Johnston, Sean A.
; REGISTRATION NUMBER: 35,910
; REFERENCE/DOCKET NUMBER: P1024
; TELEPHONE: 415/225-3562
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1575 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: DNA
US-08-639-294-1

Query Match      5.0%; Score 60.8; DB 3; Length 1575;
Best Local Similarity 44.6%; Pred. No. 4.1e-10;
Matches 395; Conservative 0; Mismatches 467; Indels 24; Gaps 3;

QY 200 AGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 259
DB 222 AGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 281
260 ACCTGATTAACAGACCAAGAGTGTCTTGGGAGAGCTTACAGATCTGTATGACACAC 319
282 CACTCATCAACAGCCCGGAGGGGCGGTGGGCGGAGGCTGCGAGCGCTGTACCGAGCA 341
320 ATAAATCCAAATGACACAGCCTATCTAATATACACGATGGTGTCCCTCGATCTGTGA 379
```









```

, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc feature
, LOCATION: (480001)..(495000)
, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc feature
, LOCATION: (495001)..(510000)
, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc feature
, LOCATION: (510001)..(525000)
, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc feature
, LOCATION: (525001)..(540000)
, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc feature
, LOCATION: (540001)..(555000)
, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc feature
, LOCATION: (555001)..(570000)
, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc feature
, LOCATION: (570001)..(585000)
, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc feature
, LOCATION: (585001)..(600000)
, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc feature
, LOCATION: (600001)..(615000)
, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc feature
, LOCATION: (615001)..(630000)
, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc feature
, LOCATION: (630001)..(645000)
, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc feature
, LOCATION: (645001)..(660000)
, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc feature
, LOCATION: (660001)..(675000)
, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc feature
, LOCATION: (675001)..(690000)
, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc feature
, LOCATION: (690001)..(705000)
, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc feature
, LOCATION: (705001)..(720000)
, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc feature
, LOCATION: (720001)..(735000)
, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc feature
, LOCATION: (735001)..(750000)
, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc feature
, LOCATION: (750001)..(765000)
, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc feature
, LOCATION: (765001)..(780000)
, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc feature
, LOCATION: (780001)..(795000)
, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc feature
, LOCATION: (795001)..(810000)
, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc feature
, LOCATION: (810001)..(825000)
, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc feature
, LOCATION: (825001)..(840000)
, OTHER INFORMATION: n=a or c or g or t

```

```

; NAME/KEY: misc_feature
; LOCATION: (840001)..(850000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (855001)..(870000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (870001)..(885000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (885001)..(900000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (900001)..(915000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature

Query Match      3.0%; Score 36.2; DB 4; Length 1230025;
Best Local Similarity 60.8%; Pred. No. 14;
Matches 59; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY      820 TGCAGAAAAGAAACACAGAGTCCTTCAACAGCTTCCCTGCCTTACCATGCTCTACAAC 879
Db      91784 TGCAGAAAATCAAATCGAGATTACATGTTCTTTTCTCTGTAATGGAGGCTCTATC 91843

QY      880 ATCAAGTCATTGGGTAACTTCCCAAGTCTTACTTCA 916
Db      91844 TTCAAGTCCTTAGATAACTCCATATACTCACAATCA 91880

RESULT 10
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid

```

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match
Best Local Similarity 2.9%; Score 36; DB 1; Length 7218;
Matches 33; Conservative 188; Mismatches 183; Indels 0; Gaps 0;

QY 35 AGAAATCACAGCAAGGCTCTTAAGAACAGTCTCTTTCTTCTCTCTCTCTCTCTCTCTCTCTCTG 94
Db 1507 AAAAAACGGCATGTAGGCATCACTGTAATTACCTATCTATGCAAGTAGTTAAAGAGATAG 1448

QY 95 GGGTCCTGGGACACACAGAAATCTCATGCAGAAATGAATATGGTCAAGCTGTGCACTGGT 154
Db 1447 AAGAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1368

QY 155 TTAICTTTTATAAGTTACCCAAAAGGACTACAGAGCAAGTGAAGAGCGGGGCTGCAGT 214
Db 1387 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1328

QY 215 ACCTGTACTCGACTCCACAACAAACCTCGAAACAAGAGCTCTACTGTATTACACGCA 274
Db 1327 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1268

QY 275 CCAGAGTGTCTCGGAGGACCTTACAGCATCTGTATGCACACATAATCCACGAATG 334
Db 1267 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1208

QY 335 ACACAGCTATCTAATATACACAGATGTTCTCTGATCTGTGAATACAGCAGACAGT 394
Db 1207 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1148

QY 395 ATGGACATGCCAAAGTCTGCTGTATGGAACAGAACGCAGGGG 438
Db 1147 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1104

RESULT 11
US-09-547-435-7/c
; Sequence 7, Application US/09547435
; Patent No. 6582957
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander, Jr.
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6582957el Lipoxigenase Proteins and Polynucleotides
; FILE OF INVENTION: Encoding the Same
; FILE REFERENCE: 7705.0009-0000
; CURRENT APPLICATION NUMBER: US/09/547,435
; CURRENT FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1005
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-547-435-7

Query Match
Best Local Similarity 2.9%; Score 35.4; DB 4; Length 1005;
Matches 75; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 629 GGAACTTATCTACATGCCCGGATGTGCCAACTCCAGTTCTCTTAAAGATCCCTGTCC 688
Db 166 GAACATTGGGCTCCATGTATACATCAGGATGGGATGTCATTTTCATGGGAAGCCGGCA 107

QY 689 GGTTACCTCGTGAATGTCACCTAGCCAGGCTCTAACTTGTGTCATTTTGCATAATCAA 748
Db 106 GGTTACCGATTCCACATGCTGTCTACCTCTGTCTACACAGTTGTGCTCTTTGTCAAGGCAA 47

```

```

QY 629 GGAACCTTATCTACATGCCCGGATGTGTCACCACTCCAGTTCCTTAAAGATCCCTGTCC 688
Db 631 GAACATTGGCTCCATGTACATCAGGATGGGATGTCATTTTCATGGGAAGCCGGCA 572
QY 689 GGTACCTCGTGAACCTGCACTCAGCCAGGCTTAAACTTCGTCCTCAATTTTGCAAAATCAA 748
Db 571 GGTACCGATTCCCACTGCTCACCCTGGTCTACACAGTTGCTGCTTTGTCAAGCAA 512
QY 749 GTTTTATCTGATGACATCT 769
Db 511 ATTCTGTCTGACTCCATCT 491

```

```

RESULT 14
US-09-547-435-9/c
; Sequence 9, Application US/09547435
; Patent No. 6582957
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander, Jr.
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6582957el Lipoxigenase Proteins and Polynucleotides
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: 7705.0009-00000
; CURRENT APPLICATION NUMBER: US/09/547,435
; CURRENT FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 1848
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-547-435-9

```

```

Query Match 2.9%; Score 35.4; DB 4; Length 1848;
Best Local Similarity 53.2%; Pred. No. 0.26;
Matches 75; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 629 GGAACCTTATCTACATGCCCGGATGTGTCACCACTCCAGTTCCTTAAAGATCCCTGTCC 688
Db 631 GAACATTGGCTCCATGTACATCAGGATGGGATGTCATTTTCATGGGAAGCCGGCA 572
QY 689 GGTACCTCGTGAACCTGCACTCAGCCAGGCTTAAACTTCGTCCTCAATTTTGCAAAATCAA 748
Db 571 GGTACCGATTCCCACTGCTCACCCTGGTCTACACAGTTGCTGCTTTGTCAAGCAA 512
QY 749 GTTTTATCTGATGACATCT 769
Db 511 ATTCTGTCTGACTCCATCT 491

```

```

RESULT 15
US-09-547-435-25/c
; Sequence 25, Application US/09547435
; Patent No. 6582957
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander, Jr.
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6582957el Lipoxigenase Proteins and Polynucleotides
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: 7705.0009-00000
; CURRENT APPLICATION NUMBER: US/09/547,435
; CURRENT FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 1938
; TYPE: DNA

```

```

; ORGANISM: Homo sapiens
US-09-547-435-25

Query Match 2.9%; Score 35.4; DB 4; Length 1938;
Best Local Similarity 53.2%; Pred. No. 0.27;
Matches 75; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 629 GGAACCTTATCTACATGCCCGGATGTGTCACCACTCCAGTTCCTTAAAGATCCCTGTCC 688
Db 1099 GAACATTGGCTCCATGTACATCAGGATGGGATGTCATTTTCATGGGAAGCCGGCA 1040
QY 689 GGTACCTCGTGAACCTGCACTCAGCCAGGCTTAAACTTCGTCCTCAATTTTGCAAAATCAA 748
Db 1039 GGTACCGATTCCCACTGCTCACCCTGGTCTACACAGTTGCTGCTTTGTCAAGCAA 980
QY 749 GTTTTATCTGATGACATCT 769
Db 979 ATTCTGTCTGACTCCATCT 959

Search completed: October 14, 2004, 05:27:48
Job time : 108.146 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2004, 00:32:26 ; Search time 465.631 Seconds  
(without alignments)  
11167.204 Million cell updates/sec

Title: US-10-790-589-1

Perfect score: 1224

Sequence: 1 tccagtcgccgcatgga.....aacgtagcgtcccaataaag 1224

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_29Jan04.\*

1: Geneseqn1980s.\*

2: Geneseqn1990s.\*

3: Geneseqn2000s.\*

4: Geneseqn2001as.\*

5: Geneseqn2001bs.\*

6: Geneseqn2002s.\*

7: Geneseqn2003as.\*

8: Geneseqn2003bs.\*

9: Geneseqn2003cs.\*

10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1224	100.0	1224	5	AAD21288 Mouse deo
2	1219.8	99.7	1652	4	AAP60715 Murine DN
3	674.4	55.1	1690	5	ABV29663 Human pro
4	674.4	55.1	1690	5	ABV23788 Human pro
5	672.8	55.0	1268	5	AAD21289 Human deo
6	653.4	53.4	1140	6	ABK92136 Prostate
7	641.8	52.4	1086	4	AAP60716 Human DNA
8	193.2	15.8	480	8	ACH19337 Human adu
9	91.6	7.5	266145	9	ADH87477 Fowlipox v
10	65	5.3	65	6	ABN55966 Mouse spl
11	60.8	5.0	1575	2	AAV02903 Human DNA
12	60.8	5.0	1915	2	AAV29137 Homo sapi
13	59.4	4.9	1201	4	ABL13359 Drosophil
14	59.4	4.9	3760	4	ABL13358 Drosophil
15	55.4	4.5	927	2	AAV29138 Bos tauru
16	53.8	4.4	65	6	ABN29670 Rat splic
17	51	4.2	2000	7	ADA71938 Rice gene
18	48.8	4.0	60	6	ABN40474 Human spl
19	48.4	4.0	2000	7	ADA71938 Rice gene
20	41	3.3	379	7	ABX42069 Bovine ES
21	39	3.2	6609	6	ABL33303 Human imm
22	36.8	3.0	2213	6	AA96558 Corn prom
23	36.6	3.0	585	4	AA144338 Probe #13

24	36.2	3.0	110000	2	AAX91990_00	Aax91990 Nucleotid
25	36	2.9	515	6	AAD29250_	Aad29250 SHH3 comp
26	36	2.9	515	6	AAD29250	Aad29250 SHH3 comp
27	36	2.9	4974	6	ABK10072	Abk10072 Plasmid p
28	36	2.9	6611	6	AAD32909	Aad32909 pBS68 pla
29	35.8	2.9	491	8	ACH30353	Ach30353 Human tes
30	35.4	2.9	1005	3	AAC61750	Aac61750 cDNA enco
31	35.4	2.9	1149	4	ABK41932	Abk41932 cDNA enco
32	35.4	2.9	1149	8	ADB59599	Adb59599 Connectiv
33	35.4	2.9	1383	3	AAC61752	Aac61752 cDNA enco
34	35.4	2.9	1470	3	AAC61748	Aac61748 cDNA enco
35	35.4	2.9	1848	3	AAC61751	Aac61751 cDNA enco
36	35.4	2.9	1938	3	AAC61759	Aac61759 cDNA enco
37	35.4	2.9	2136	6	ABA96003	Aba96003 Human lip
38	35.4	2.9	2236	3	AAC61749	Aac61749 cDNA enco
39	35.4	2.9	2307	6	ABZ11921	Abz11921 Human pol
40	35.4	2.9	2316	3	AAC61760	Aac61760 cDNA enco
41	35.4	2.9	2604	3	AAC61758	Aac61758 cDNA enco
42	35.4	2.9	2701	3	AAC61747	Aac61747 cDNA enco
43	35.4	2.9	3320	6	ABA05868	Aba05868 Human lip
44	35.4	2.9	3320	9	AAD60556	Aad60556 Human lip
45	35.4	2.9	3384	3	AAC61761	Aac61761 cDNA enco

## ALIGNMENTS

RESULT 1	
AAD21288	
ID	AAD21288 standard; cDNA; 1224 BP.
XX	
AC	AAD21288;
XX	
DT	28-JAN-2002 (first entry)
XX	
DE	Mouse deoxyribonuclease (DNase) II beta cDNA.
XX	
KW	Mouse; deoxyribonuclease; DNase II beta protein; mucous plug; lung;
KW	Cystic fibrosis; pulmonary; ss.
XX	
OS	Mus sp.
XX	
FH	Key
CDS	Location/Qualifiers
FT	40..1104
FT	/*tag= a
FT	/product= "Mouse deoxyribonuclease (DNase) II beta
FT	protein"
XX	
PN	WO200175082-A1.
XX	
PD	11-OCT-2001.
XX	
PF	02-APR-2001; 2001WO-US010635.
XX	
PR	03-APR-2000; 2000US-00541840.
PR	19-MAY-2000; 2000US-00574942.
XX	
PA	(DART-) DARTMOUTH COLLEGE.
XX	
PI	Eastman AR, Krieser RJ;
XX	
DR	WPI; 2001-662972/76.
DR	P-FSDB; AAE13014.
XX	
PT	New cDNA encoding a deoxyribonuclease II beta enzyme useful for degrading
PT	DNA present in the mucous plugs in the lungs of cystic fibrosis patients.
XX	
PS	Claim 2; Page 13-14; 21pp; English.
XX	
CC	The invention relates to deoxyribonuclease (DNase) II beta proteins and
CC	their corresponding cDNAs. The DNase II beta may be useful to digest DNA
CC	in the mucous plugs in lungs of cystic fibrosis patients and so reduce
CC	their viscosity. The present sequence is mouse DNase II beta cDNA

```

XX SQ Sequence 1224 BP; 351 A; 308 C; 256 G; 309 T; 0 U; 0 Other;
Query Match 100.0%; Score 1224; DB 5; Length 1224;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCAGTCCCTGCATGGAATGAAGGCCACAGATAGAAATGACAGCAAGCCCTTAAGA 60
DB 1 TCCAGTCCCTGCATGGAATGAAGGCCACAGATAGAAATGACAGCAAGCCCTTAAGA 60
QY 61 ACAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 120
DB 61 ACAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 120
QY 121 TGCAGAAATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
DB 121 TGCAGAAATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 181 ACTAGCAAGGCAAGTGAAGAGGGGGCTGCAGTACCTGCTGCTGCTGCTGCTGCTG 240
DB 181 ACTAGCAAGGCAAGTGAAGAGGGGGCTGCAGTACCTGCTGCTGCTGCTGCTGCTG 240
QY 241 ACCTGGAACCAAGCCCTTACCTGATTAACAGCACCAAGGAGTCTCTGGGAGGACCTTA 300
DB 241 ACCTGGAACCAAGCCCTTACCTGATTAACAGCACCAAGGAGTCTCTGGGAGGACCTTA 300
QY 301 CAGCATCTGATGACACATATATCCAGATGACACAGGATGATGATGATGATGATGATGAT 360
DB 301 CAGCATCTGATGACACATATATCCAGATGACACAGGATGATGATGATGATGATGATGAT 360
QY 361 GGTGTCCCTGGATCTGTGAATACAGCAGACAGATGATGATGATGATGATGATGATGAT 420
DB 361 GGTGTCCCTGGATCTGTGAATACAGCAGACAGATGATGATGATGATGATGATGATGAT 420
QY 421 TGGAAACAGACGCGAGGGGCTTCTGCTGATACACTCTGTTCCCAAGTTTCCCCAGTTTCAT 480
DB 421 TGGAAACAGACGCGAGGGGCTTCTGCTGATACACTCTGTTCCCAAGTTTCCCCAGTTTCAT 480
QY 481 GGCTATGATGACCAACCTCGGGGAGGATGATGATGATGATGATGATGATGATGATGATG 540
DB 481 GGCTATGATGACCAACCTCGGGGAGGATGATGATGATGATGATGATGATGATGATGATG 540
QY 541 GGATACAGCAGGTTGAGGAATAGATTTTCAGCTCTTGTCTTACACCAACCAATCTAC 600
DB 541 GGATACAGCAGGTTGAGGAATAGATTTTCAGCTCTTGTCTTACACCAACCAATCTAC 600
QY 601 AGCTGCTTCATTCACAGCACCTTTTCACTGGAACCTTATCTACATGCCCCGGATGTGCC 660
DB 601 AGCTGCTTCATTCACAGCACCTTTTCACTGGAACCTTATCTACATGCCCCGGATGTGCC 660
QY 661 AACTCCAGTTCCTTAAAGATCCCTGTCCGGTACCTCGTGAACCTGCACTGACCCAGGTT 720
DB 661 AACTCCAGTTCCTTAAAGATCCCTGTCCGGTACCTCGTGAACCTGCACTGACCCAGGTT 720
QY 721 CTAACTTCGTCCTTTTGAATCAAGTTTCTTACTGATGATGATGATGATGATGATGATGATG 780
DB 721 CTAACTTCGTCCTTTTGAATCAAGTTTCTTACTGATGATGATGATGATGATGATGATGATG 780
QY 781 ATAGCTCAAAAGTTGAAGACACATTTGTAGCACAACCTGCGCAAAAAGAAACAAGAG 840
DB 781 ATAGCTCAAAAGTTGAAGACACATTTGTAGCACAACCTGCGCAAAAAGAAACAAGAG 840
QY 841 CTTCCTTCAAACTGTTCCTGCTTACCATGTCTCAACATCAAGTCCATGGGGTAACT 900
DB 841 CTTCCTTCAAACTGTTCCTGCTTACCATGTCTCAACATCAAGTCCATGGGGTAACT 900
QY 901 TCCAGTCTTACTTCAAGTCTTCGCCAGACCAATTCCAAATGTGTGTTTCCATAAAGGCG 960
DB 901 TCCAGTCTTACTTCAAGTCTTCGCCAGACCAATTCCAAATGTGTGTTTCCATAAAGGCG 960
QY 961 TCCGCAAAATCGCTGCACTGCTATGAGACCTTAATGCAAGCCTTACACCAAGCCTTAAGA 1020

```

```

DB 961 TCCGCAAAATCGCTGGACCTGCAATTTGAGACCTTAAATCGAAGCCTCACCAAGCCTTAAGA 1020
QY 1021 GGTGGAGGATTCATCTGTACAAAGAAATCACTACATTTACAGGCAATTCATTAATTAAT 1080
DB 1021 GGTGGAGGATTCATCTGTACAAAGAAATCACTACATTTACAGGCAATTCATTAATTAAT 1080
QY 1081 CTCGGTTATGGGTTCTGTAAAGTAAACTCGGTGAAAGGCCACACCTCTGCTTGAAGAAC 1140
DB 1081 CTCGGTTATGGGTTCTGTAAAGTAAACTCGGTGAAAGGCCACACCTCTGCTTGAAGAAC 1140
QY 1141 ACTGGCAGCTGGAACATCTGCTTGGATCTGTTCTCCATAATATCAAGGCTTCTGAGTGA 1200
DB 1141 ACTGGCAGCTGGAACATCTGCTTGGATCTGTTCTCCATAATATCAAGGCTTCTGAGTGA 1200
QY 1201 GCACAACTAGCTGCTCAATAAAG 1224
DB 1201 GCACAACTAGCTGCTCAATAAAG 1224

RESULT 2
AAF60715
ID AAF60715 standard; cDNA; 1652 BP.
XX AAF60715;
AC AC
DT 03-MAY-2001 (first entry)
XX Murine DNase coding sequence.
DE Murine; DNase; DLAD; deoxyribonuclease; endonuclease; cystic fibrosis;
KW infectious disease; ss.
XX Mus musculus.
XX WO200112793-A1.
PN 22-FEB-2001.
PD 01-MAY-2000; 2000WO-JP002893.
PP 17-AUG-1999; 99JP-00230870.
PR (TANU/) TANUMA S.
PA Tanuma S, Shiohawa D;
PI WPI; 2001-218348/22.
XX P-PSDB; AAB72416.
DR Acidic deoxyribonuclease capable of divalent cation-independent cleavage
of DNA under acidic even neutral pH and not inhibited by G-actin, useful
in remedies for cystic fibrosis and for treatment of infectious diseases.
XX Claim 14; Page 49-51; 61pp; Japanese.
XX The present sequence is the coding sequence of a murine deoxyribonuclease
(DLAD), which is an endonuclease. DLAD is capable of divalent cation-
independent cleavage of DNA under acidic conditions. DLAD can be used as
a substitute for DNase I in treating cystic fibrosis, and is useful in
the prevention and treatment of infectious diseases
XX Sequence 1652 BP; 494 A; 410 C; 339 G; 409 T; 0 U; 0 Other;
SQ Query Match 99.7%; Score 1219.8; DB 4; Length 1652;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1221; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCCAGTCCCTGCTGCAATGAAGGCCACAGATAGAAATGACAGCAAGCCTCTAAGAA 61
DB 175 CACAGTCCCTGCTGCAATGAAGGCCACAGATAGAAATGACAGCAAGCCTCTAAGAA 234
QY 62 CAGTCTTCTTCTGCTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121

```



Db	235	CAGTCTCTTTCTTGTGCTCTCTTTGGCCCTCTCTGGGGTCTCTGGGGAACACAGAAAATCTCAT	294
Qy	122	GCAGAAATGAATATGCTGAAGCTGTGGACTGTGTTATCTTTTATAGTTTACCCAAAGGA	181
Db	295	GCAGAAATGAATATGCTGAAGCTGTGGACTGTGTTATCTTTTATAGTTTACCCAAAGGA	354
Qy	182	CTAGCAAGCAAGTAGTGAAGGGGGGGCTGCAGTACTGTACTTGGACTCCACAAGACAA	241
Db	355	CTAGCAAGCAAGTAGTGAAGGGGGGGCTGCAGTACTGTACTTGGACTCCACAAGACAA	414
Qy	242	CCTGGAACAAGAGCCCTTACCTGATTTACAGCACACAGAGTGCTCTGGGGAGGACCTTAC	301
Db	415	CCTGGAACAAGAGCCCTTACCTGATTTACAGCACACAGAGTGCTCTGGGGAGGAGCCTTAC	474
Qy	302	AGCATCTGTATGACACACATAAATCCAGAAATGACACAGCCCTATCTTAATATACAAGATG	361
Db	475	AGCATCTGTATGACACACATAAATCCAGAAATGACACAGCCCTATCTTAATATACAAGATG	534
Qy	362	GTGTCCCTGGATCTGTGAATTTACGACAGACAGTATGGACATGCCAAAGGTCTGCTGGTAT	422
Db	535	GTGTCCCTGGATCTGTGAATTTACGACAGACAGTATGGACATGCCAAAGGTCTGCTGGTAT	594
Qy	422	GGACAGAAACGACGGGTTCTGGCTGATACACTCTGTTCCCAAGTTTCCCCAGTTCTAG	481
Db	595	GGACAGAAACGACGGGTTCTGGCTGATACACTCTGTTCCCAAGTTTCCCCAGTTCTAG	654
Qy	482	GCTATGAGTACCCAACTCGGGGAGCGGATATGGACAAACCGGACTGTGCATCATCTTTCG	541
Db	655	GCTATGAGTACCCAACTCGGGGAGCGGATATGGACAAACCGGACTGTGCATCATCTTTCG	714
Qy	542	GATACAGCGAGTTTGAGAAATAGATTTTCAGTCTTGGTCTTACAACCAACACTCTACA	601
Db	715	GATACAGCGAGTTTGAGAAATAGATTTTCAGTCTTGGTCTTACAACCAACACTCTACA	774
Qy	602	GCTGCTTCATTTCCAAAGCACTTTCACCTGGAACTTATCTACATGCCCGGATGTGTGCCA	661
Db	775	GCTGCTTCATTTCCAAAGCACTTTCACCTGGAACTTATCTACATGCCCGGATGTGTGCCA	834
Qy	662	ACTCCAGTTCCTTAAAGATCCCTGTCCGGTACCTCGCTGGAAGTGCACCTCAGCCACAGGTC	721
Db	835	ACTCCAGTTCCTTAAAGATCCCTGTCCGGTACCTCGCTGGAAGTGCACCTCAGCCACAGGTC	894
Qy	722	TAAACTTCGTCCATTTTGCAAAATCAAAGTTTATCTGATGACATCTTTACAGATGGA	781
Db	895	TAAACTTCGTCCATTTTGCAAAATCAAAGTTTATCTGATGACATCTTTACAGATGGA	954
Qy	782	TAGCTCAAAAGTTCAAGACACATTTGTTAGACACAAACCTGGCAGAAAGAAACAAGAGC	841
Db	955	TAGCTCAAAAGTTCAAGACACATTTGTTAGACACAAACCTGGCAGAAAGAAACAAGAGC	1014
Qy	842	TTCCCTTCAAACTGTTCCTCGCTTACCATGTCTACAACATCAAGTCCATTTGGGGTAACTT	901
Db	1015	TTCCCTTCAAACTGTTCCTCGCTTACCATGTCTACAACATCAAGTCCATTTGGGGTAACTT	1074
Qy	902	CCAAAGCTTATCTTCAGTTCTCGCAAGACCATTCCTCAAAATGTTGTGTTTCCATAAAGGCT	961
Db	1075	CCAAAGCTTATCTTCAGTTCTCGCAAGACCATTCCTCAAAATGTTGTGTTTCCATAAAGGCT	1134
Qy	962	CCCAAAATCCGTGGACCTGTGATTTGGAGACCTTAATTCGAAGCCTTACACAAAGCCTTAAGAG	1021
Db	1135	CCGCAAAATCCGTGGACCTGTGATTTGGAGACCTTAATTCGAAGCCTTACACAAAGCCTTAAGAG	1194
Qy	1022	GTGGAGGATTCATCTGTACAAAGAACTCACTACATTTTACCGGACATTTCAATAATTATATC	1081
Db	1195	GTGGAGGATTCATCTGTACAAAGAACTCACTACATTTTACCGGACATTTCAATAATTATATC	1254
Qy	1082	TCCGTTATGGGTTCTGTAAAGTAAACTCGGTGAAGGCGCACACCTCTGTCTCTTGAAGAAACA	1141
Db	1255	TCCGTTATGGGTTCTGTAAAGTAAACTCGGTGAAGGCGCACACCTCTGTCTCTTGAAGAAACA	1314
Qy	1142	CTGGCACTGAAACATCTCGCCTTGGATCTGTTCTTCATAATATCAAGGCTTCTGAGTGAG	1201
Db	1315	CTGGCACTGAAACATCTCGCCTTGGATCTGTTCTTCATAATATCAAGGCTTCTGAGTGAG	1374

QY	1202	CACAACGTAGCGTCCATAAAAG	1224
Db	1375	CACAACGTAGCGTCCATAAAAG	1397

### RESULT 3

ABV29663  
ID ABV29663 standard; cDNA; 1690 BP.

AC ABV29663;

XX  
DT 16-SEP-2002 (first entry)

Human prostate expression marker cDNA 29654.

Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW KW pharmacogenomic marker; gene; ss.

OS Homo sapiens.

PN WO200160860-A2.

AA PD 23-AUG-2001

XX  
PF 20-FEB-2001-

17 FEB 2000 0000 UTC

PR 16-MAR-2000; 2000US-0189862P.

PR 09-JUN-2000; 2000US-0211314P.

PR 13-DEC-2000; 2000US-0255281P.

PA (MILL-) MILLENNIUM PREDICTIVE

PI Schlegel R, Endege WO, Monahan JE;

DR WPI; 2001-662795/76.

Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

PS Claim 1; Page 6369; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) also useful as a pharmacodynamic or pharmacogenomic marker

SQ Sequence 1690 BP; 525 A; 369 C; 305 G; 485 T; 0 U; 6 Other:

Query Match 55.1%; Score 674.4; DB 5: Length 1690;

test local similarity 75.0%; Pred. No. 4.2e-206;  
Matches 872; Conservative 0; Mismatches 281;  
Indels 9; Gaps 2.

QY  
31 AGATAGAAATGACAGCAAGCCCTCTAAGAACAGTTCCTTTGGTCTTCTTCCCCC

Dd  
Dd

37 AAACAGAAAATGATTGGCAAGAATTGCTAAGAACATCCTTTGCTTTCCTCTCCCTCCCCCCC

QY 91 TCTGGGGTCCCTGGGGACACCAGAAATCTCATGCAGAAATGAATAATGCTGAAGCTCTCCAC 150

db  
97 TTTGGGGTGTCTGGGGGACGCAACAATTTTCATGCACGAATCTACACACCGCAATACCTCTCCAC

151 TGGTTTATCTTTTATAGTACCCAAAGGACTAGCAAGCAAGTGAAGAGCGGCGCTG 210  
157 TGGTTTACTTTTATAGTACTTAAAGACAAACAAAGAAAGTGGAGAGCTGGGTTA 216  
211 CAGTACTGTACTGACTCCACAGACAAACCTCGGAAACAAAGAGCCTTACCTGATTAA 270  
217 GAGTACTGTACTGACTCCACAGACAAACCTCGGAAACAAAGAGTGGAGAGTGAAT 276  
271 AGCAGAGAGTGTCTCGGAGAGCCTTACAGATCTGTATGACACACATAATTCACG 330  
277 GACACCAAGAGTCTTTTGGAGAGCATTACACAGCTATATCAAGCATATGCTCTAAG 336  
331 AATG---ACACGCTTCTTAATATACAGAGTGTCTCGGAAACAAAGAGTGGAGAGT 387  
337 AGTAAACAACACGCTTCTTAATATACAGAGTGTCTCGGAAACAAAGAGTGGAGAGT 396  
388 AGCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 447  
397 AGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 456  
448 ATACACTGTCTTCCAAAGTTCCTCCAGTTC-----ATGGCTATGATGATGATGATGAT 501  
457 ATTCAATCCATCCCTCAGTTCCTCCAAATTCGGAGAGAGGCTATGATATCCACCCACA 516  
502 GGGAGGGATATGACACAAACCGGATCTGATCACTTTGATACAGCCAGTTCGAGGAA 561  
517 GGGAGAGCAATGACAAAGTGGATCTGATCACTTTGATACAGCCAGTTCGAGGAA 576  
562 ATAGATTTTTCAGTCTTGGTCTTCAACACCAACATCTACAGCTGCTTCATTCCAGCACC 621  
577 ATAGATTTTTCAGTCTTGGTCTTCAACACCAACATCTACAGCTGCTTCATTCCAGCACC 636  
622 TTTCATGGAACCTTATCTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 681  
637 TTTCATGGAACCTTATCTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 696  
682 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 741  
697 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 756  
742 AAATCAAGTCTTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 801  
757 AAGTGGATCTTCTTCTGACGATCTTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 816  
802 CATTGTTAGACAAACCTGCGAGAAAGAAACAGAGCTTCTTCAAACTGTCCTG 861  
817 CACTTGTATACAGAAACCTGCGAGAAAGAAACAGAGCTTCTTCAAACTGTCCTG 876  
862 CTTTACCATGCTACACATCAAGTCCATTGGGTTAATCTTCAAGTCTTACTTCAGTCT 921  
877 CTTTACCATGCTACACATCAAGTCCATTGGGTTAATCTTCAAGTCTTACTTCAGTCT 936  
922 CGCAAGACCATTCCTCAATGGTGTGTTTCCATAAGGCTCCGCAATTCGCTGAGCTGTC 981  
937 TATCAAGTATCATGCAAGTGTGTTTCTCCAAAGGGGCAACAAATTCGCTGAGCATGT 996  
982 ATTGAGACCTTAATCGAAGCTTACCAAGCTTACAGAGTGGAGGATTCATCTGTACA 1041  
997 ATTGAGACCTTAATCGAAGCTTACCAAGCTTACAGAGTGGAGGATTCATCTGTACC 1056  
1042 AAGTATCACTATTTTACAGGCTATTCATTAATTTATATCTCCGTTATGGGTTCTGTAA 1101  
1057 CAGATTTGGCAATTTTACAGCATTTTCAAGTATGATTTATATCTATGAAAGCTGTAA 1116  
1102 TAACTCGGTGAAAGGCCACACCTCTGCTCTTGTGAAACACATGCGACCTGGAACATCTCG 1161  
1117 TAACTCGGTGAAAGGCCACACCTCTGCTCTTGTGAAACACATGCGACCTGGAACATCTCG 1176  
1162 CTTGATCTGTTCTCCATAA 1183  
1177 ATTACACCTTCTTTATATTTA 1198

RESULT 4  
ABV23788  
ID ABV23788 standard; cDNA; 1690 BP.  
XX  
AC ABV23788;  
XX  
DT 16-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 23779.  
XX  
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
OS Homo sapiens.  
XX  
XX WO200160860-A2.  
PN  
XX 23-AUG-2001.  
XX  
XX 20-FEB-2001; 2001WO-US005171.  
XX  
XX 17-FEB-2000; 2000US-0183319P.  
PR 16-MAR-2000; 2000US-0189862P.  
PR 25-MAY-2000; 2000US-0207454P.  
PR 09-JUN-2000; 2000US-0211314P.  
PR 18-JUL-2000; 2000US-0219007P.  
PR 13-DEC-2000; 2000US-0255281P.  
XX  
(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
PA  
XX Schlegel R, Endege WO, Monahan JE;  
PI  
XX WPI; 2001-662795/76.  
XX  
XX Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
XX Claim 1; Page 4376; 11750pp; English.  
XX  
XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
SQ Sequence 1690 BP; 525 A; 369 C; 305 G; 485 T; 0 U; 6 Other;  
Query Match 55.1%; Score 674.4; DB 5; Length 1690;  
Best Local Similarity 75.0%; Pred. No. 4.2e-206;  
Matches 872; Conservative 0; Mismatches 281; Indels 9; Gaps 2;  
QY 31 AGATGAGAAATGACAGCAAGCCCTTAAGACAGTCTTCTTCTTCTTCTTCTTCTTCTTCT 90  
Db 37 AAACAGAAATGATGGCAAGACTGCTAAGAACATCTTCTTCTTCTTCTTCTTCTTCTTCT 96  
QY 91 TCTGGGTCCTGGGACACACCAATCTCATGCAAAATGAATGTTGTAAGCTGTGGAC 150  
Db 97 TTTGGGTCCTGGGACACCAATCTCATGCAAAATGAATGTTGTAAGCTGTGGAC 156  
QY 151 TGGTTTATCTTTTATAGTTCACCAAGAGTCTAGCAGGCAAGTGAAGAGCGGGCTG 210  
Db 157 TGGTTTATCTTTTATAGTTCACCAAGAGTCTAGCAGGCAAGTGAAGAGCTGGTTA 216  
QY 211 CAGTACTGTACTGACTCCACAGACAAACCTCGGAAACAAAGAGTGGAGAGTGAAT 270



Db 397 AGAAGATGGAACACCAAGGTTTACTGCTGTGGAACAGAGTTCAAGGTTCTGGCTG 456  
Qy 448 ATACACTCTGTTCCCAAGTTTCCCGAGTTC-----ATGGCTATGAGTACCCAACTCG 501  
Db 457 ATTCAATCCATCCTCAGTTTCTCTCAATTCGGGAAGAGGCTATGATTATCCACCCACA 516  
Qy 502 GGAGGCGATATGGAACAAACCGGCTCTGATCAGTCTTGGATACAGCCAGTTTGAGGA 561  
Db 517 GGAGACGAAATGGACAAAGTGGCATCTGCATAACTTTCAGGTACCAACAGATGAGGCA 576  
Qy 562 ATAGATTTTCACTCTTGTGCTTTACACCAACCAATCTACAGCTGCTTCATTCCAAGCACC 621  
Db 577 ATAGATTTTCACTCTTGTGCTTTACACCAACCAATCTACAGCTGCTTCATTCCAAGCACC 636  
Qy 622 TTTCACTGGAACCTTATCTACATGCCCCCGGATGTGCGCAACTCCAGTTCCTTAAAGATC 681  
Db 637 TTTCAACGAGGCTCAATTCATGCCCCCAGCTGTGCACAGGCGCCAGCTCATCAGAGATT 696  
Qy 682 CTTGTCGGTACCTCGTGAATGCTGACTCAGCCAGGCTTAACCTCGTCCATTTGCA 741  
Db 697 CTTGGCAGGCTCTCTCAACACACTTCAGTCGGCCGAGGACAAAAATTCCTCCATTTGCA 756  
Qy 742 AAATCAAGTTTATATCTGATGACATCTTTACAGATGGATAGTCTCAAAAAGTTGAAGACA 801  
Db 757 AAGTCGGATCTTTCTTGAGCACTCTTTGCGCCCTGGATGCTCAACGGCTGAAGACA 816  
Qy 802 CATTTGTTAGCACAAACCTGGCAGAAAAGAAACAGAGCTTCTTCAAACTGTTCCCTG 861  
Db 817 CACTTTGTTAAACAGAAACCTGGCAGGAAAAGAAACAGAGCTTCTTCAAACTGTTCCCTT 876  
Qy 862 CTTTACATGCTTACACATCAAGTCCATTTGGGTAATCTCCAGTCTTACTTCACTTCT 921  
Db 877 CTTTACATGCTTACATATAAAGCAATTAATATACGACACTCTTATTTCACTTCT 936  
Qy 922 CGCAACAGCAATTCCTAATGTTGTTTCCATAAGGGCTCCGCAAAATCGTGACCTGC 981  
Db 937 TATCAAGATCAGCCAGTGGTGTATTTCCAAAGGGCCACCAAAATCGTGACATGT 996  
Qy 982 ATTGGAGACCTTAATGAAGCCTACACCAAGCCTTAAGAGGTGAGGATTCATCTGACA 1041  
Db 997 ATTGGAGACCTTAATGGAGTCCACACCAAGCCTTCAGAGTGGAGGATTCATTTGACC 1056  
Qy 1042 AAGATCAGTACATTTACCGGCAATTCATAAATATATCTCCGTATGGTCTCTGAAG 1101  
Db 1057 CAGAAATGGCAAAATTTACCAAGCAATTTCAAGGATTAGTATTAATTAAGCTGTAAG 1116  
Qy 1102 TAAACTCGGTGAAGGCCACACCCCTCTGCTTCTGAAACACTGGCAGCTGGAAACATCTGC 1161  
Db 1117 TAACTTGGTGAAGGACACAGAGTACTATCATTTGAAACCTTGACAAATGGGTCTTCTCC 1176  
Qy 1162 CTTGGATCTGTTCTCCATAATA 1183  
Db 1177 ATTACACCTTCTTTATATTTTA 1198

RESULT 6  
ID ABK92136  
XX standard; DNA; 1140 BP.

XX AC ABK92136;

XX DT 15-AUG-2002 (first entry)

XX Prostate cancer-associated DNA sequence #22.

XX prostate cancer; prostate tumour tissue; human; mammal; cytostatic;  
XX gene therapy; gene; ds.

XX Mammalia.

XX WO200230268-A2.

PD 18-APR-2002.  
XX 12-OCT-2001; 2001WO-US032045.  
XX 13-OCT-2000; 2000US-00687576.  
PR 08-DEC-2000; 2000US-0073288.  
PR 08-DEC-2000; 2000US-00733742.  
PR 24-JAN-2001; 2001US-0263957P.  
PR 16-MAR-2001; 2001US-0276791P.  
PR 16-MAR-2001; 2001US-0276888P.  
PR 06-APR-2001; 2001US-0281922P.  
PR 24-APR-2001; 2001US-0286214P.  
PR 30-APR-2001; 2001US-00847046.  
PR 04-MAY-2001; 2001US-0288589P.  
XX (BOSB-) EOS BIOTECHNOLOGY INC.  
PA Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;  
XX P-PSDB; ABG61821.  
XX WPI; 2002-471335/50.

DR Detecting a prostate cancer-associated transcript in a cell in a patient,  
DR useful for diagnosing prostate cancer (PC) or screening modulators of PC,  
DR by determining if prostate cancer-associated genes are expressed in a  
DR prostate tissue.

XX Claim 22; Page 317; 436pp; English.

PS The present invention relates to methods of detecting a prostate cancer-  
XX associated transcript in a cell from a patient. The method comprises  
XX contacting a biological sample from the patient with prostate cancer-  
XX associated polynucleotides (designated PC genes) that selectively  
XX hybridise to a sequence that is at least 80% identical to them. The  
XX prostate cancer-associated polynucleotide sequences are differentially  
XX expressed in prostate tumour tissue or in prostate cancer and are derived  
XX from the tissues of various organisms such as humans or other mammals  
XX (e.g. mice, sheep and dogs). The methods of the invention are useful for  
XX diagnosing and treating prostate cancer in mammals. The prostate cancer-  
XX associated genes are useful for diagnosing or treating prostate cancer,  
XX as well as for identifying modulators of prostate cancer or agents that  
XX inhibit prostate cancer. The nucleic acid sequences are particularly  
XX useful in gene therapy, as a vaccine or in antisense applications.  
XX ABK92115-ABK92263 represent prostate cancer-associated polynucleotide  
XX sequences

XX Sequence 1140 BP; 357 A; 258 C; 245 G; 280 T; 0 U; 0 Other;

SQ Query Match 53.4%; Score 653.4; DB 6; Length 1140;  
Best Local Similarity 75.9%; Pred. No. 2e-199;  
Matches 836; Conservative 0; Mismatches 256; Indels 9; Gaps 2;

Qy 31 AGATAGAAATGACAGCAAGCCTTAAGACAGTCTTTCTTGTCTCTTTGGCTCTCTTTGGCCCTC 90  
Db 37 AACAGAAATGATGGCAAGACTGCTAAGACATCTTTTGTCTCTTTCTTTCTTTCTTTGGCCCTC 96  
Qy 91 TCTGGGTCTCTGGGACACCAACCAATCTCATGCAAAATGAATATGTTGAAGCTGTGGAC 150  
Db 97 TTTGGGTCTCTGGGACGACCAATTTTATGCAAAATGAAGAGGAAAGCTGTGGAC 156  
Qy 151 TGGTTATCTTTTATAGTTACCCAAAGGACTAGCAAGGCAAGTGAAGAGCGGGCTG 210  
Db 157 TGGTTTACTTTTATAGTTTACCTTAAAGACAAACCAAGAAAGTGGAGACTGGGTTA 216  
Qy 211 CAGTACCTGTACTGGACTCCACAGACAAACCTTGGAAACAGAGCCTTACTCTGATTAAC 270  
Db 217 GAGTACCTGTACTAGACTCTTACACTAGAGCTGGAGAAAGAGTGAAGCTAATGAAT 276  
Qy 271 AGCACCAGGAGTCTCTGGGAGGACCTTACAGCATCTGTATGACACATTAATCCAG 330  
Db 277 GACACCAAGAGTGTGTTTGGGAAGGACATTACACAGCTATATGAAGCATATGCTCTTAAG 336  
Qy 331 ---AATGACACAGCCTATCTAATATACACGATGTTCCCTGGATCTGTGATTAACAGC 387



Db 604 TTTCACGAGCTCATTACATGCCCCAGCTGTGACACGAGCCAGCTCATCAGAGATT 663

QY 682 CCTGTCGGTACCTCGTGAAGTGAAGTCTGAGCCAGGGTCTAACTTCGTCATTTTGA 741

Db 664 CTTGCGAGGCTCTCACCACCTTCAGTCCGCGCCAGGACCAAAATTCCTCCATTTTGA 723

QY 742 AATCAAGTTTATATCTGATGATGATCTTTACAGGATGGATGCTCAAAAGTTGAAGACA 801

Db 724 AAGTGGATTTCTTTCTTGATGGCATCTTTGCGAGCTGGATGGCTCAACGGCTGAAGACA 783

QY 802 CATTTCTGTAGCACAACTCGGACAGAAAAAGAAACAGAGCTTCCTTCAAACTGTTCCCTG 861

Db 784 CATTGTTTACAGAACTCGGACGGAAGAAAGACAGAGCTTCCTTCAAACTGCTCCCTT 843

QY 862 CTTTACATGCTTACCAATCAAGTCAATTTGGGGTAACTTCCAGTCTTACCTTCAGTTCT 921

Db 844 CTTTACATGCTTACCAATCAAGTCAATTTGGGGTAACTTCCAGTCTTACCTTCAGTTCT 903

QY 922 CGCAGAGACCATTCCTCAATGCTGTGTTTCCATAAAGGCTCCGCAATCGCTGACCTGC 981

Db 904 TATCAAGATCATGCCAAGTGGTGTATTTCCAAAGGGACCAAAATCGCTGACATGT 963

QY 982 ATTGGAGACCTTAATCGAAGCTTACACCAAGCCCTTAAGAGTGGAGGATTCATCTGTACA 1041

Db 964 ATTGGAGACCTTAATCGAAGCTTACACCAAGCCCTTAAGAGTGGAGGATTCATCTGTACC 1023

QY 1042 AAGAACTACATCAATTTACCGAGGCAATTTCAATAATTAATCTCCGTTATGGTCTGTAAAG 1101

Db 1024 CAGAAATGGCAAAATTTACCAAGCAATTTCAAGGATTAGTATTATATATGAAGCTGTAAAG 1083

QY 1102 TAA 1104

Db 1084 TAA 1086

RESULT 8

ACH19937

ID ACH19937 standard; cdna; 480 BP.

AC ACH19937;

DT 13-OCT-2003 (first entry)

XX Human adult lung cdna #940.

XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;

KW genome mapping; biodiversity; genetic disorder.

XX Homo sapiens.

XX US2003073623-A1.

PD 17-APR-2003.

XX 30-JUL-2001; 2001US-00918995.

PF 30-JUL-2001; 2001US-00918995.

XX 30-JUL-2001; 2001US-00918995.

XX (DRMA/) DRMANAC R T.

PA (LABA/) LABAT I.

PA (STAC/) STACHE-CHRAIN B.

PA (DICK/) DICKSON M C.

PA (JONE/) JONES L W.

XX Drmanac RT, Labat I, Stache-Chrain B, Dickson MC, Jones LW;

XX WPI; 2003-615964/58.

XX New polynucleotide sequences obtained from various cdna libraries, useful

PT as hybridization probes, as oligomers for PCR, for chromosome and gene

PT mapping, in the recombinant production of protein, or in generating

PT antisense DNA or RNA.

XX

PS Claim 1; SEQ ID NO 7149; 44pp; English.

XX The invention relates to an isolated polynucleotide comprising any one of

CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was

CC determined by the technique of SBH (sequencing by hybridisation). Also

CC included is a purified polypeptide comprising a sequence corresponding to

CC a reading frame of the novel polynucleotide. The nucleic acid sequences

CC are useful in diagnostics as expressed sequence tags (EST) for

CC identifying expressed genes or for physical mapping of the human genome,

CC in forensics, in assessing biodiversity, or in identifying mutations

CC responsible for genetic disorders and other traits. The nucleotide

CC sequences are also useful as hybridisation probes, as oligomers for PCR,

CC for chromosome and gene mapping, in the recombinant production of

CC protein, or in generating antisense DNA or RNA. The purified polypeptide

CC is useful for generating antibodies specific for it. The present sequence

CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data

CC for this patent did not form part of the printed specification, but was

CC obtained in electronic format directly from USPTO at

CC seqdata.uspto.gov/sequence.html?DocID=20030073623

XX

SQ Sequence 480 BP; 134 A; 120 C; 106 G; 118 T; 0 U; 2 Other;

Query Match 15.8%; Score 193.2; DB 8; Length 480;

Best Local Similarity 79.2%; Pred. No. 3e-51;

Matches 244; Conservative 0; Mismatches 58; Indels 6; Gaps 1;

QY 331 AATGACACAGCCTATCTAATATATACACGATGGTGTCCCTGGATCTGTGAATTACAGAGA 390

Db 171 AACACACAGCCTATCTAATATATACATGATGGATGCTCTAAACCTGTGAATTACAGAGA 230

QY 391 CAGTATGGACATGCCAAGGTCGTGGTATGGAACAGAACGCGAGGGTTCGTGCTGATA 450

Db 231 AAGTATGGACACCAAGGTTTACTGCTGTGGAACAGAGTTCAAGGGTTCGTGCTGATT 290

QY 451 CACTCTGTTCCAGATTTCCCGCAGTTC-----ATGCGTATAGTACCCAACTCCGGG 504

Db 291 CATTCCATCCCTCAGTTTCTCTCCAAATTCGGAAGAGGCTATGATATCCACCACAGGG 350

QY 505 AGCGATATGACAAACCGGCATCTGCATCACTTCGCGATACAGCCAGTTTGAGGAATA 564

Db 351 AGAGAAATGGACAAAGTGGCATCTGCATACTTTCAAGTACACACCATGATGAGCAATA 410

QY 565 GATTTTCAGCTCTTGGTCTTACAAACCAACATCTACAGCTGCTTCATTCCAGCACCTTT 624

Db 411 GATTCTCAGCTCTTGGTCTGCAACCCCAACGCTCTATAGCTGCTCCATCCAGGCACCTTT 470

QY 625 CACTGGAA 632

Db 471 CACCAGGA 478

RESULT 9

ADE87477/C

ID ADE87477 standard; DNA; 266145 BP.

XX ADE87477;

AC ADE87477;

DT 29-JAN-2004 (first entry)

XX Fowlpox virus genome DNA.

XX fowlpox virus; FPV; virucide; tuberculostatic; protozoacide; antipyrctic;

XX cytotatic; hepatotropic; antibacterial; vaccine; malaria; tuberculosis;

KW East Coast fever; avipox virus; influenza; hepatitis;

KW human papilloma virus; tumour; leishmaniasis; listeriosis; theileria;

KW gene; ds.

XX Fowlpox virus.

OS WO2003047617-A2.

XX 12-JUN-2003.

XX

PF 02-DEC-2002; 2002WO-GB005411.  
 XX  
 PR 30-NOV-2001; 2001GB-00028733.  
 PR 30-NOV-2001; 2001US-0334649P.  
 XX  
 PA (ISIS-) ISIS INNOVATION LTD.  
 XX  
 PI Laidlaw S, Skinner M, Hill A, Gilbert S, Anderson R;  
 XX  
 DR WPI; 2003-513700/48.  
 XX  
 XX Treating and/or preventing e.g. malaria or tuberculosis, or eliciting an  
 PT immune response, comprises administering a priming composition and a  
 PT boosting composition containing a non-replicating viral vector in either  
 PT order.  
 XX  
 XX Claim 30; SEQ ID NO 1; 302pp; English.  
 PS  
 XX The invention relates to a fowlpox virus (FPV) genome which has  
 CC modifications in one or more wild-type FPV genes. The invention further  
 CC relates to a novel method for treating and/or preventing a disease in a  
 CC subject comprising administering two compositions, each containing a non-  
 CC replicating viral vector. At least one of the compositions comprises a non-  
 CC poxvirus vector derived from a fowlpox virus. The novel compositions have  
 CC the following activities: virucide, tuberculostatic, protozoacide,  
 CC antipyretic, cytostatic, hepatotropic, and antibacterial. The non-  
 CC replicating viral vector is useful in a vaccine for an animal.  
 CC particularly a mammal such as a primate, specifically human. The non-  
 CC medicament for treating and/or preventing a disease which is, or results  
 CC from, a chronic infection such as malaria, tuberculosis or East Coast  
 CC fever, or for eliciting a T-cell immune response in a subject. Non-  
 CC cultured CEF cells are useful for growing an avipox virus, such as  
 CC fowlpox virus. The method or the vaccine may further be used to treat or  
 CC prevent influenza, hepatitis, human papilloma virus and other viral  
 CC infections, malignancies such as tumours, leishmaniasis, listeriosis, and  
 CC theileria. This polynucleotide sequence represents the DNA of the fowlpox  
 CC virus genome of the invention.  
 XX  
 SQ Sequence 266145 BP; 92375 A; 41026 C; 40846 G; 91897 T; 0 U; 1 Other;  
 Query Match 7.5%; Score 91.6; DB 9; Length 266145;  
 Best Local Similarity 49.8%; Pred. No. 6.4e-17;  
 Matches 474; Conservative 0; Mismatches 439; Indels 39; Gaps 8;  
 QY 121 TGCAGAAATGAATATGCTGAAGCTGGAGCTGTTTATCTTTATAGTTACCCAAAGG 180  
 DB |||||  
 QY 38771 TGTGTAATGAAGAGGCGAAGTAGTAGTGGTATTTCTGTATACAGCTCCCTAAGCTC 38712  
 DB |||||  
 QY 181 ACTAGCAAGGCAAGTGAAGAGCGGGCTGCTAGTACCTGCTGACTCCCAAGACAA 240  
 DB |||||  
 QY 38711 CAAAAATTAGGCACTAAA-----GTAATGAATATTATATACATAGATGTAATCT 38658  
 DB |||||  
 QY 241 ACCTGGACACAGAGCTCTACTGCTGATTAACACACAGAGTGCTCTGGGGAGGACCTTA 300  
 DB |||||  
 QY 38657 AAATCGAAACCGAGGTAAGTAGTACCTTAACTTAGATACAGTATTAATAGGGAACACTA 38598  
 DB |||||  
 QY 301 CAGCATCTGTATGACACACATATTTCCAGATGACACAGCCTATCTAATATACAGCAT 360  
 DB |||||  
 QY 38597 TATCTATTATGACTATATGAT---AGTAATATATAGATAATCATTTTATATATGAT 38541  
 DB |||||  
 QY 361 GGTGTCCTGTGATCTGTGAATATACAGCAGACAGTATGACATGCCAAAGTCTGCTGCTA 420  
 DB |||||  
 QY 38540 GGCATCCAGGAAGCAAAATTTATTCATGAGAGGTAGTCACTAATAGGCTTATGCA 38481  
 DB |||||  
 QY 421 TGGACAGAACCCAG-----GGGTTCTGGCTGTATACACTCTGTTCCCAAGTTTCCCCCA 474  
 DB |||||  
 QY 38480 TGGATATGATAGCGTAAACAGGATTTTGGTTGATACATCTCTGTACCTAGATTCCCACT 38421  
 DB |||||  
 QY 475 GTTCAAT-----GGCTATGATPACCAACCTCGGAGGCGATATGGCAACCGGATC 528  
 DB |||||  
 QY 38420 TCTCTCTGTGTGGGATACAAATATCTTACTCTGCTGTATGATACGGGAATCTATGTTA 38361  
 DB |||||

QY 529 TGCATCACATTTGGGATACA---GCCAGTTTGAGGAATAGATTTTACGCTCTTGGTCTTA 585  
 DB |||||  
 QY 38360 TGTATTAACCTTGGATTACAAAGGAGGTCTTACGGCCTTAGACATAACACTACCTGTTAAT 38301  
 DB |||||  
 QY 586 CAACCAAAACATCTACAGCTGCTTCAITCCCAAGCACCTTTTCACTGGAAACTTATCTACATG 645  
 DB |||||  
 QY 38300 AACCCAAACGTGTATAACTGTTCAGTAACAATAAGAACTGTAATATCTATACCATCTA 38241  
 DB |||||  
 QY 646 CCCGGATGTGTCGCAACTCCAGTTCCTTAAAGATCCCTGTCGGTACCTCGGTGAACATG 705  
 DB |||||  
 QY 38240 TGTAAACGATAAAACATATACGACTTTATATAAAATGTTTCTAGGTGGATGG----- 38189  
 DB |||||  
 QY 706 CACTCAGCCAGAGGTCTAAACTTCGTCCTTTTGGCAAAATCAAGTTTATATCTGATGAC 765  
 DB |||||  
 QY 38188 -AATCTAGAAAGGAGGAGAAAGTTTAAACATTTGCTAAGTCAAAATATTTTCCCATGAT 38130  
 DB |||||  
 QY 766 ATCTTTACAGGATGATAGCTCAAAAGTTGAAGACACATTTGTTAGCACAAAACCTGGCAG 825  
 DB |||||  
 QY 38129 ATAATGAGTGCAATGATAGGTCCGACTTTGGAATCTGCACTATACAGAAACATGGCAA 38070  
 DB |||||  
 QY 826 AAAAAAGAAACAGAGCTTCCCTCAAACTGTTCCCTGCCCTTACCATGTTTACAACTCAAG 885  
 DB |||||  
 QY 38069 CBTAGAGGAGAGATGATGATAACCAATTCGTCGCAAAATACCATGTACATAACATAAAA 38010  
 DB |||||  
 QY 886 TCCATTGGGGTAACTTCCCAAGTCTTACTTTCAGTTCTCGCCCAAGACCATTCCAATGGTGT 945  
 DB |||||  
 QY 38009 TCGATAAACGTAATGGAACATCGTTTATTAATTAAT---ATGATCATTTCTAAATGGATT 37953  
 DB |||||  
 QY 946 GTTTCATTAAGGGCTCCGCAATCGCTGAGCTGATTTGGAGACCTTAATCGAAGCCTA 1005  
 DB |||||  
 QY 37952 GTTTCATTTACGATAA---AAAAAGATGGGTATGTATAGTGTATATAGGATCTCT 37896  
 DB |||||  
 QY 1006 CACCAAGCCTTAAGAGGTGGAGATTCATCTGTACAAAGAACATCACTACATTT 1057  
 DB |||||  
 QY 37895 ACTCAACGTCATAGAGGAGGTGGATACGCTGTACTAGGAATGGATATCTGT 37844  
 DB |||||  
 RESULT 10  
 AEN55966  
 ID AEN55966 standard; DNA; 65 BP.  
 XX  
 AC AEN55966;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:28714.  
 XX  
 KW Human; mouse; rat; splice transcript; detection; RNA transcript;  
 KW splice variant; transcriptome; oligonucleotide library; ss.  
 XX  
 OS Mus musculus.  
 XX  
 FN WO200210449-A2.  
 XX  
 PD 07-FEB-2002.  
 XX  
 PF 20-JUL-2001; 2001WO-IB001903.  
 XX  
 PR 28-JUL-2000; 2000US-0221607P.  
 PR 02-MAY-2001; 2001US-0287724P.  
 XX  
 XX (COMP-) COMPUEN INC.  
 PA  
 PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;  
 XX  
 XX WPI; 2002-257383/30.  
 DR  
 PT New oligonucleotide libraries comprising oligonucleotides which  
 PT selectively hybridize to mRNAs transcribed from a transcription unit of a  
 PT genome, useful for detecting tissue-, pathology-, and developmental-  
 PT specific genes.  
 XX  
 PS Example 1; SEQ ID NO 28714; 47pp; English.



XX The present invention describes oligonucleotide libraries for detecting  
 CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-  
 CC )transcriptome comprises messenger RNAs transcribed from multiple  
 CC transcription units that populate a genome. The library comprises several  
 CC oligonucleotides, each capable of hybridizing selectively to a set of  
 CC messenger RNAs transcribed from a given transcription unit of the genome,  
 CC which encodes one or more messenger RNA splice variants. The  
 CC oligonucleotide libraries are useful for detecting mRNAs from a  
 CC biological sample, in expression profiling studies, in qualitatively or  
 CC quantitatively characterizing the corresponding transcriptome, and in  
 CC detecting RNA transcripts and splice variants of human or animal  
 CC transcriptomes. The libraries may also be used as specialised mini-  
 CC libraries to detect transcripts of a sub-transcriptome under a particular  
 CC biological or pathological state, and so allowing the detection of tissue  
 CC - and pathology-specific genes such as those genes only expressed in  
 CC specific tissue under a specific pathological condition; to detect  
 CC developmental specific genes; and to detect RNA transcripts and splice  
 CC variants of a transcriptome of a patient suffering from a particular  
 CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from  
 CC rats, humans and mice, which are used in the exemplification of the  
 CC present invention. N.B. The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 65 BP; 15 A; 22 C; 10 G; 18 T; 0 U; 0 Other;  
 SQ  
 Query Match 5.3%; Score 65; DB 6; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-10;  
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1118 CCACACCTCTGCTTGAACACTGCGACTGGAACATCTGCGCTGGATCTGTTCTCC 1177  
 DB 1 CCACACCTCTGCTTGAACACTGCGACTGGAACATCTGCGCTGGATCTGTTCTCC 60  
 QY 1178 ATAT 1182  
 DB 61 ATAT 65  
 RESULT 11  
 AAV02903  
 ID AAV02903 standard; DNA; 1575 BP.  
 AC AAV02903;  
 XX 08-MAY-1998 (first entry)  
 DT Human DNase II DNA.  
 DE  
 XX DNase II; human; deoxyribonuclease; pulmonary disease; diagnostic;  
 KW cellular DNA debris; ss.  
 XX Homo sapiens.  
 XX  
 PH Key Location/Qualifiers  
 FT misc\_feature 29..31  
 FT /tag= a  
 FT /note= "Region indicated in specification"  
 FT CDS 86..1168  
 FT /tag= b  
 FT /product= "DNase II"  
 FT sig\_peptide 86..133  
 FT /tag= c  
 FT mat\_peptide 134..1165  
 FT /tag= d  
 FT /product= "DNase II"  
 XX WO9740134-A2.  
 PN  
 XX 30-OCT-1997.  
 PD  
 XX 23-APR-1997; 97WO-US006664.  
 PF

XX 25-APR-1996; 96US-00639294.  
 PR (GETH ) GEMENTECH INC.  
 XX Baker KP, Baron WF;  
 PI WPI; 1997-535820/49.  
 XX P-PSDB; AAW31495.  
 DR Human DNase II and related nucleic acids - useful in protein production  
 XX e.g. for therapeutic use to treat systemic lupus erythematosus and  
 PT pulmonary diseases e.g. cystic fibrosis.  
 PS Claim 1; Fig 1; 30pp; English.  
 XX This sequence encodes a novel human deoxyribonuclease, DNase II. This  
 CC protein is useful to reduce the viscoelasticity of DNA-containing  
 CC material, e.g. mucus, and used to treat patients with pulmonary diseases  
 CC or disorders e.g. cystic fibrosis, or with systemic lupus erythematosus.  
 CC It may also improve antibiotic efficacy in the treatment of abscesses,  
 CC infected lesions etc. provide treatment in non-infected conditions in  
 CC which there is an accumulation of cellular DNA debris, e.g.  
 CC pyelonephritis and be used to degrade DNA in biological samples or in  
 CC diagnostic assays. The nucleic acids and vectors can be used for in vivo  
 CC or ex vivo gene therapy, and antisense oligonucleotides can be produced  
 CC from the nucleic acids which can bind to and prevent expression of  
 CC nucleic acid within cells. The nucleic acids (or a portion) can also be  
 CC used for hybridisation assays for nucleic acids encoding human DNase II  
 CC in a sample, or to identify and isolate nucleic acids sharing substantial  
 CC sequence identity (e.g. encoding naturally-occurring allelic variants of  
 CC human DNase II). The antibodies can be used to detect and measure human  
 CC DNase II in tissues or clinical samples, and in the purification of human  
 CC DNase II. The nucleic acids enable production of human DNase II by  
 CC recombinant DNA methods in quantities sufficient for clinical use, not  
 CC previously possible  
 XX Sequence 1575 BP; 358 A; 465 C; 450 G; 302 T; 0 U; 0 Other;  
 SQ  
 Query Match 5.0%; Score 60.8; DB 2; Length 1575;  
 Best Local Similarity 44.6%; Pred. No. 2.9e-08;  
 Matches 395; Conservative 0; Mismatches 467; Indels 24; Gaps 3;  
 QY 200 AGCGGGGCTGCAGTACTGTACTGACTGCAACAGCAAACTGGAACAGAGCTCT 259  
 DB 222 AGAGAGGCTGCAGTACTGCAAGTATCTGACGAGAGCTCGGAGGCTGGCGGACGCGCAG 281  
 QY 260 ACCTGATTAACAGCACGAGGAGTCTCTGGGAGGAGCTTACAGCATCTGTATGACACAC 319  
 DB 282 CACTCATCAACAGCCCGGAGGGGCGTGGGCGGAGCTGACGCGCTGTACCGGAGCA 341  
 QY 320 ATATTTCCAGATGACAGACCTTATCTATATACAGATGGTGTCTCCTGGATCTGTGA 379  
 DB 342 ACACAGGAGCTGCTCTCTCTCTACATGACCAACCCCTCAACCCAGAGGCTC 401  
 QY 380 ATTACAGCAGACAGTATGACATCCAAAGTCTGTGGTATGGAACAGAGCGGGGT 439  
 DB 402 AGGACTCTTCATCGCTGGGACACAGAGGCTGTCTCTCTTACACAGATGGGGCT 461  
 QY 440 TCTGGTGATACACTCTGTTCCCAAGTTTCCCGAGTTTCATGGTATGAGTACCAACT 499  
 DB 462 TCTGGTGGTCCACAGTGATACCTAATCTCCCTCCACCGGCTCTCTCTGCTGATACAGCT 521  
 QY 500 CG-----GGGAGGCGATATGGACAAACCGGCATCTGCATCACTTTCGGATACAGCC 550  
 DB 522 GGCCTCATAGCGCTGTACCTACGGGACAGCCCTCTGTGTCTCTTCCCTCGCTC 581  
 QY 551 AGTTTGAGGAATAGATTTTTCAGCTCTTGGTCTGTACACCAACATCTACAGCTCTTCA 610  
 DB 582 AGTTCTCGAAGATGGGCAAGCAGCTGACCTACACCTACCCCTGGGTCTATAACTACAGC 641  
 QY 611 TTCCAGCAGCTTTTACCTGGAACCTTATCTACATGCCCGGAGTGTGTCCAACTCCAGTT 670



Db 642 TGGAGGGATCTTTGGCCAGGAATTCGCCAGCTTGGAGATGTGGTCAAGGGCCACCAAG 701  
 QY 671 CCTTAAAGATCCCTGTCCGGTACCTGCTGAATGCACTCAGCCAGGCTTAAACTTCG 730  
 Db 702 TTAGCCAGAACCTGGAACAGAGCATCACATCCAGCCGGGGCTGTTTTC 761  
 QY 731 TCCATTTGCAAAATCAAGTTTTTATCTGATGACATCTTTACAGGATGATAGCTCAAA 790  
 Db 762 AGAGCTTTGCCAAGTTCCAGCAATTTGGAGATGACCTGTACTCCGGCTGTTGGCAGCAG 821  
 QY 791 AGTTGAAGACACATTTGTAGCACAACTGGCAGAAAAGAAACAAGACTTCTTCAA 850  
 Db 822 CCTTGGTCAACCTGCAGGTCAGTTCTGGCACAAAACCTGTAGGCATCTCGCCTCTA 881  
 QY 851 ACTGTTCCCTGCCTTACCATGCTACAAATCAAGTCCATTGGGT-----AACTTCCA 904  
 Db 882 ACTGCTCGGATATCTGGCAGTTCTGAATGGAACAGATAGCTTCCCTGGACCCAGCG 941  
 QY 905 AGTCTTACTCAGTTCTGGCAAGACCATTTCCAAATGGTGTGTTTCCATAAAGGGCTCCG 964  
 Db 942 GCCCAAGCTTCAACAGCAGAGGACCACTCCAAATGGTGTGCTGCCCAAAAGG----- 996  
 QY 965 CAAATCGCTGGACCTGCATTTGAGCACTAAATCGAAGCTTACACAGCTTAAAGAGTG 1024  
 Db 997 ----GCCCTGGACCTGGTGGGTGACATGAATCGGAACCGAGGAGAGCAACGGGGTG 1052  
 QY 1025 GAGGATTCATCTGTACAAAGAAATCACTATATTACAGGCAATTCA 1070  
 Db 1053 GGGGCACACTGTGTGCCAGCTGCCAGCCTCTGGAAAGCCTTCCA 1098

## RESULT 12

AAV29137  
 ID AAV29137 standard; cDNA; 1915 BP.  
 AC AAV29137;  
 XX  
 DT 11-SEP-1998 (first entry)  
 DE Homo sapiens DNase II gene.  
 XX  
 KW DNase II; deoxyribonuclease II; cystic fibrosis; DNA digestion;  
 KW controlled cell death; apoptosis; metapneumonia; cell turnover;  
 KW neurodegeneration; AIDS; acquired immune deficiency syndrome; cancer;  
 KW autoimmune disease; diagnosis; anticancer agents;  
 KW enhanced chromosomal rearrangement; chromosome instability; ss.  
 OS Homo sapiens.  
 XX  
 FH Key  
 FT CDS  
 FT 68..1150  
 FT /\*tag= a  
 FT /product= "deoxyribonuclease (DNase II)"  
 XX  
 FN W09816659-A1.  
 XX  
 PD 23-APR-1998.  
 XX  
 PF 09-OCT-1997; 97WO-US018262.  
 XX  
 PR 15-OCT-1996; 96US-0028539P.  
 XX  
 PA (DART-) DARTMOUTH COLLEGE.  
 XX  
 PI Eastman A, Krieser R;  
 XX  
 DR WPI; 1998-251301/22.  
 XX  
 DR P-PSDB; AAW37920.  
 XX  
 PT Human and bovine deoxyribonuclease II enzyme and encoding cDNA - useful,  
 PT e.g. to induce apoptosis in tumour cells, identify agents modulating  
 PT apoptosis and digest DNA in sputum of cystic fibrosis patients.  
 XX

PS Claim 2; Page 15-16; 29pp; English.

XX The sequence is that encoding deoxyribonuclease II (DNase II). The enzyme  
 CC is useful to digest DNA, e.g. in the lung sputum of cystic fibrosis  
 CC patients to reduce sputum viscosity. It (or fragments) are also useful to  
 CC produce antibodies, e.g. to study DNase II expression in cells. DNase II  
 CC has recently been linked with DNA fragmentation in the early stages of  
 CC controlled cell death (apoptosis), a process critical to homeostasis  
 CC during, e.g. metamorphosis or cell turnover. Too much cell death can lead  
 CC to neuro-degeneration and acquired immune deficiency syndrome (AIDS),  
 CC whilst too little can lead to cancer or autoimmune diseases. The  
 CC antibodies raised against it can therefore be used to diagnose apoptotic  
 CC stages in selected cells, by contacting cells with the antibody,  
 CC detecting binding of the antibody with DNase II and determining DNase II  
 CC levels. Such diagnosis is useful to evaluate the efficacy of therapeutic  
 CC agents, e.g. anticancer agents to promote apoptosis in cells. The  
 CC antibodies can also be used to identify cells susceptible to premature  
 CC death. The cDNA is useful to identify agents modulating apoptosis in  
 CC cells, by treating cells with an agent, transfecting cells with cDNA and  
 CC monitoring apoptosis compared with untreated cells. Inhibitors identified  
 CC may be useful in preventing diseases relating to enhanced chromosomal re-  
 CC arrangement. Vectors comprising the cDNA can be used to induce apoptosis  
 CC in selected cells, e.g. tumour cells or cells involved in autoimmune  
 CC disorders. Antisense oligonucleotides can be administered to cells to  
 CC inhibit DNase II expression to reduce chromosome instability associated  
 CC with cancer  
 XX

Sequence 1915 BP; 462 A; 533 C; 530 G; 390 T; 0 U; 0 Other;

Query Match 5.0%; Score 60.8; DB 2; Length 1915;  
 Best Local Similarity 44.6%; Pred. No. 3.3e-08;  
 Matches 395; Conservative 0; Mismatches 467; Indels 24; Gaps 3;  
 QY 200 AGCGGGGCTGCAGTACCTGTACTGGACTCCACAGACAAACCTCGAAGAGCTCT 259  
 Db 204 AGAGAGGGCTGCAGTACAGTATCTGGACAGAGCTCCGGAGGCTGGCGGCGGAGG 263  
 QY 260 ACCTGATTAACAGACAGCAGGAGTGTCTTGGGAGGACCTTACAGATCTGTATGACAC 319  
 Db 264 CACTCATCAACAGCCCGAGGGGGCGGTGGGCGGAGCTGCGAGCGCTGTATCCGAGCA 323  
 QY 320 ATAATTTCCAGGAATGACAGCAGCTATCTAATATACAGATGTTCTCCCTGGATCTGTA 379  
 Db 324 ACACAGCAGCTCCCTCTCTGCTCTACATGACCAACCGCTCAACCCAGCAAGGCTC 383  
 QY 380 ATTACAGCAGACAGTATGACATGCGAAGGTCTCTGTGTATGGAACAGAACGCGAGGT 439  
 Db 384 AGGACTCTTCCATGCTGGGCACAGAGGGTGTCTGTCTTGCACCATGAGTGGGGT 443  
 QY 440 TCTGGCTGATACACTCTGTTCCTCAAGTTTCCCGAGT-----TCATGGCTATGAGT 490  
 Db 444 TCTGGCTGGTCCACAGTGTACTTAACCTTCCCTCCACCGGCTCTCTGTGATACAGTT 503  
 QY 491 ACCCAACCTCGGGGAGCGGATATGACAAACCGGCATCTGCATCACTTTCGGATACAGCC 550  
 Db 504 GGCCTCATAGCGCTGTACTACGGGAGAGCCCTGCTGTGTGTCTTTCCTTCCGCTC 563  
 QY 551 AGTTTGAGGAATAGATTTTTCAGCTCTTGTGTTTACAAACCAACATCTACAGCTGTTCA 610  
 Db 564 AGTTCTCGAAGATGGGCAAGCAGCTGACCTACACCTACCCCTGGGTCTATACTACCAGC 623  
 QY 611 TTCCNAGCAGCTTTTCACTGGAAACTTATCTACATGCCCCGGATGTGTGCAACTCCAGTT 670  
 Db 624 TGGAGGGGATCTTTGCCCGAGGAATTTCCCGACTTGGAGATGTGTGTCAGGCGCCACG 683  
 QY 671 CCTTAAAGATCCCTGTCGGGTACCTCGCTGACCTGACCTCAGCCAGGCTTAACCTCG 730  
 Db 684 TTAGCCAGAACCTCGAAGCAGAGCATCACTACATCCAGCCCGGGGCTGTTTTCC 743  
 QY 731 TCCATTTTGCAAAATCAAGTTTTTATCTGATGATCTTTTACAGGATGGATAGTCAAA 790  
 Db 744 AGAGCTTTGCCAAGTTCCAGCAAAATTTGGAGATGACCTCTACTCCGGCTGTGTGGCAGCAG 803

QY 791 AGTTGAGACACATTTGTAGCACAACTGGCAGAAAGAAAGAACAGAGCTTCCTTCAA 850  
 Db 804 CCGTTGGTACCAACCTGCGAGTCCAGTTCTGGCACAATACTGAGCATCTCCCTCTTA 863  
 QY 851 ACTGTTCCCTGCTTACCATGCTCAACATCAAGTCCCAATGCGGT-----AACTTCCA 904  
 Db 864 ACTGCTCGATATCTGGCAGTTCTGATGTGACCAAGATAGCTTCCCTGGACAGCGG 923  
 QY 905 AGTCTTACTTCAGTTCTGCGCAAGACCATTCCTCAAAATGGTGTGTTTCCATAAAGGGCTCCG 964  
 Db 924 GCCCAAGCTTCAACAGCACAGAGGACCATCTCCAAATGGTGTGTTCCCAAAAGG----- 978  
 QY 965 CAAATCGCTGGACCTGATTTGAGACCTAAATCGAAGCCTACACCAAGCCTTAAGAGGTG 1024  
 Db 979 -----GCCCTGGACCTGCTGGGTGACATGATCGGAACCGAGGAGGACCAAGGGTG 1034  
 QY 1025 GAGGATTCATCTGACAAAATCACTACATTTACCAGGCAATTTCA 1070  
 Db 1035 GGGGCACACTGTGTGCCAGCTGCCAGCCCTCTGGAAAGCCTTCCA 1080

RESULT 13  
 ABL13359  
 ID ABL13359 standard; cDNA; 1201 BP.  
 AC ABL13359;

26-MAR-2002 (first entry)  
 Drosophila melanogaster expressed polynucleotide SEQ ID NO 34559.  
 Drosophila; developmental biology; cell signalling; insecticide;  
 pharmaceutical; gene; ss.  
 Drosophila melanogaster.

WO200171042-A2.  
 27-SEP-2001.  
 23-MAR-2001; 2001WO-US009231.  
 23-MAR-2000; 2000US-0191637P.  
 11-JUL-2000; 2000US-00614150.  
 (PEKE ) PE CORP NY.  
 Venter JC, Adams M, Li PWD, Myers EW;  
 WPI; 2001-656860/75.  
 P-PSDB; ABB69256.  
 New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.

Claim 1; SEQ ID NO 34559; 21bp + Sequence Listing; English.  
 The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511); expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 Query Match 4.9%; Score 59.4; DB 4; Length 1201;  
 Best Local Similarity 55.6%; Pred. No. 7.1e-08;

Matches 114; Conservative 0; Mismatches 91; Indels 0; Gaps 0;  
 QY 396 TGGACATGCCAAGGTCCTGCTATGGAACAGACGAGGGGTTCTGCTGATACACTC 455  
 Db 366 TGGTCATGCCAAGGAGTGTGGCCAGTGTGGGAGACGGCTATTGATTCTCCACTC 425  
 QY 456 TGTTCCTCAAGTTTCCCTCCAGTTTCATGGCTATGATACCCCAACCTCGGGAGGCGATANGG 515  
 Db 426 GTGCCCCCAATTTCCCAACAATTCGGGACTACAGCTATCCACATCCGGCGAGCAGTACGC 485  
 QY 516 ACAAACCGGCATCTGTCATCCTTTCCGATACAGCCAGTTTGAGGAAATAGATTTCAGCT 575  
 Db 486 ACAGAGCATGCTCTGGCTCACACTGAAGGCGAGGATTTGGAGAAGTTGTCAGATTCT 545  
 QY 576 CTGTGCTTTACACCAACCAATCTAC 600  
 Db 546 GGTGTACATGAACCGCATTCTAC 570

RESULT 14  
 ABL13358  
 ID ABL13358 standard; cDNA; 3760 BP.  
 AC ABL13358;

26-MAR-2002 (first entry)  
 Drosophila melanogaster expressed polynucleotide SEQ ID NO 34556.  
 Drosophila; developmental biology; cell signalling; insecticide;  
 pharmaceutical; gene; ss.  
 Drosophila melanogaster.

WO200171042-A2.  
 27-SEP-2001.  
 23-MAR-2001; 2001WO-US009231.  
 23-MAR-2000; 2000US-0191637P.  
 11-JUL-2000; 2000US-00614150.  
 (PEKE ) PE CORP NY.  
 Venter JC, Adams M, Li PWD, Myers EW;  
 WPI; 2001-656860/75.  
 P-PSDB; ABB69255.  
 New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.  
 Claim 1; SEQ ID NO 34556; 21bp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511); expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 Query Match 4.9%; Score 59.4; DB 4; Length 3760;  
 Best Local Similarity 55.6%; Pred. No. 1.4e-07;  
 Matches 114; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 396 TGGACATGCCAAAGGTCTGCTGATGGAACAGAACAGCGGGTCTCGGTGATACACTC 455  
 |||||  
 Db 1925 TGGTCATGCCAAAGAGTGTGGCCAGTGTGAGGAGCGCTATTGGATTGCTCACTC 1984  
 |||||  
 QY 456 TGTTCCTCAAGTTCCTCCAGTTCATGCTATGAGTACCCAACTTCGCGGAGGATAGG 515  
 |||||  
 Db 1985 GGTGCCCAATTTCCAAATTCGGACTACAGCTATCCACATCCCGGAGCAGTACGC 2044  
 |||||  
 QY 516 ACAACCGGCATCTGCATCACTTTCCGATACAGCCAGTTTGAGGAATAGATTTCAGCT 575  
 |||||  
 Db 2045 ACAGAGATGCTCTGCTGCTACACTGAAGGCGGAGATTGGAGAAGTTGGTCAATTTCT 2104  
 |||||  
 QY 576 CTTGGTCTTACAAACCAACATCTAC 600  
 |||||  
 Db 2105 GGTGTACATGAACCGCACTTCTAC 2129  
 |||||

## RESULT 15

AAV29138  
 ID AAV29138 standard; cDNA; 927 BP.

XX AC  
 XX AAV29138;  
 XX

DT 11-SEP-1998 (first entry)

XX Bos taurus DNase II gene.

DE DNase II; deoxyribonuclease II; cystic fibrosis; DNA digestion;  
 KW controlled cell death; apoptosis; metaplasia; cell turnover;  
 KW neurodegeneration; AIDS; acquired immune deficiency syndrome; cancer;  
 KW autoimmune disease; diagnosis; anticancer agents;  
 KW enhanced chromosomal rearrangement; chromosome instability; ss.

XX Bos taurus.

XX Key Location/Qualifiers

FT I. 831

FT /\*tag= a

FT /product= "deoxyribonuclease II (DNase II)"

XX WO9816659-A1.

XX 23-APR-1998.

XX 09-OCT-1997; 97WO-US018262.

XX 15-OCT-1996; 96US-0028539P.

XX (DART-) DARTMOUTH COLLEGE.

XX Eastman A, Krieser R;

XX WPI; 1998-251301/22.

XX P-PSDB; AAW37921.

PT Human and bovine deoxyribonuclease II enzyme and encoding cDNA - useful,  
 PT e.g. to induce apoptosis in tumour cells, identify agents modulating  
 PT apoptosis and digest DNA in sputum of cystic fibrosis patients.

PS Claim 2; Page 17; 29pp; English.

XX The sequence is that encoding deoxyribonuclease II (DNase II). The enzyme  
 CC is useful to digest DNA, e.g. in the lung sputum of cystic fibrosis  
 CC patients to reduce sputum viscosity. It (or fragments) are also useful to  
 CC produce antibodies, e.g. to study DNase II expression in cells. DNase II  
 CC has recently been linked with DNA fragmentation in the early stages of  
 CC controlled cell death (apoptosis), a process critical to homeostasis  
 CC during, e.g. metaplasia or cell turnover. Too much cell death can lead  
 CC to neuro-degeneration and acquired immune deficiency syndrome (AIDS),  
 CC whilst too little can lead to cancer or autoimmune diseases. The  
 CC antibodies raised against it can therefore be used to diagnose apoptotic  
 CC stages in selected cells, by contacting cells with the antibody,  
 CC detecting binding of the antibody with DNase II and determining DNase II

CC levels. Such diagnosis is useful to evaluate the efficacy of therapeutic  
 CC agents, e.g. anticancer agents to promote apoptosis in cells. The  
 CC antibodies can also be used to identify cells susceptible to premature  
 CC death. The cDNA is useful to identify agents modulating apoptosis in  
 CC cells, by treating cells with an agent, transfecting cells with cDNA and  
 CC monitoring apoptosis compared with untreated cells. Inhibitors identified  
 CC may be useful in preventing diseases relating to enhanced chromosomal re-  
 CC arrangement. Vectors comprising the cDNA can be used to induce apoptosis  
 CC in selected cells, e.g. tumour cells or cells involved in autoimmune  
 CC disorders. Antisense oligonucleotides can be administered to cells to  
 CC inhibit DNase II expression to reduce chromosome instability associated  
 CC with cancer  
 XX

SQ Sequence 927 BP; 214 A; 258 C; 237 G; 218 T; 0 U; 0 Other;

Query Match 4.5%; Score 55.4; DB 2; Length 927;

Best Local Similarity 46.3%; Pred. No. 1.2e-06;

Matches 311; Conservative 0; Mismatches 336; Indels 24; Gaps 3;

QY 396 TGGACATGCCAAAGGTCTGCTGATGGAACAGAACAGCGGGTCTCGGTGATACACTC 455

|||||  
 Db 75 TGGGCACACGAGGGTGTGCTCTCTGGACCAAGAGGGGGCTTCTGGTTGATCCACAG 134

|||||  
 QY 456 TGTTCCTCAAGTTCCTCCAGTTCATGCTATGAGTACCCAACTTCG 506

|||||  
 Db 135 CGTTCCTCAAGTTCCTCCAGTTCATGCTATGAGTACCCAACTTCG 194

|||||  
 QY 507 GCGATATGGACAAACCGGATCTGCATCACTTTCCGATACAGCCAGTTTGAGAAATAGA 566

|||||  
 Db 195 AAAATATGGGACACCTGATCTGTATCTTTCTCTCACCACATCTCTGGATATCAG 254

|||||  
 QY 567 TTTTCAGCTCTGGTCTTACAAACCAACATCTACAGCTGCTTCATTCGAGCACTTTCA 626

|||||  
 Db 255 CAACACAGCTGACCTATACCTATCTGATATGACACAGCTGGAGGGGACTTTGG 314

|||||  
 QY 627 CTGGAACCTTATCTACATGCCCGGATGTGTCACACTCCAGTTCCTTAAAGATCCCTGT 686

|||||  
 Db 315 CCAGAAATTCCTCTACCTCGAGGAGGTAGTCAAGGCGCATCAGTTCCGCGAGGACCGTG 374

|||||  
 QY 687 CCGGTACCTCGCTGAACCTGCACTCAGCCCGGCTCTAAACTTCGTCTCAATTTTGCAAAATC 746

|||||  
 Db 375 GAACAGCAGTGTAACTCACTCACAATCAAGAAAGAGGCCACATTCAGAGCTTTGCCAATTT 434

|||||  
 QY 747 AAGTTTATATCTGATGATCATCTTTACAGGATGATAGTCTAAAAGTTGAGACACATTT 806

|||||  
 Db 435 TGGAAACTTTGGAGATGACCTGTACTCTGCTGGTGGCGAAGCCCTTGGCAGTACCCCT 494

|||||  
 QY 807 GTTAGCACAAACCTCGCGAGAAAGAAACAGAGCTTCCTTCAAACTGTTCCTGCTTTA 866

|||||  
 Db 495 GCAGGTCCAAATTCGCGAACGATCTTCTGGATTCCTTGGCTTCAACTCTGCTCTGGGGCCCA 554

|||||  
 QY 867 CCATGCTTACACATCAAGTCCATTTGGGGTAACT-----TCCAAAGTCTTACTTCAAGTTC 920

|||||  
 Db 555 GCATGTTTGGACGTGACTCAGACAGCTTTCCCTGGCGCAGCTGGCGCAGCTTCAATGC 614

|||||  
 QY 921 TCGCCCAAGACCATTCCAAATGTTGTTTCCATAAAGGCTCCGCAAAATTCGTGACCTG 980

|||||  
 Db 615 CACAGAAAGCCATTCCAAAGTGTGTAAACCCCAAAAGG-----GCCCTGGGCGCTG 665

|||||  
 QY 981 CATTGGAGACCTTAAATCGAAGCCTTACCAAGCCTTTAAGAGGTGAGGATTCATCTGTAC 1040

|||||  
 Db 666 TGTGGGTGACATGAATCGGAACCAAGAGAGAGCACCAGGGGTGGGGGCACTCTGTGTGC 725

|||||  
 QY 1041 AAGAATCACT 1051

|||||  
 Db 726 CCAGATGCTCT 736

Search completed: October 14, 2004, 01:02:01  
 Job time : 470.631 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2004, 00:33:41 ; Search time 4727.53 Seconds  
(without alignments)  
11221.906 Million cell updates/sec

Title: US-10-790-589-1  
Perfect score: 1224  
Sequence: 1 tccagtcctcgtcatgaa.....aacgtagctccaataaag 1224

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_man.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	1224	100.0	1224	6	AR201143	AR201143 Sequence
2	1219.8	99.7	1652	6	AR432046	AR432046 Sequence
3	1219.8	99.7	1652	6	BD091714	BD091714 Novel deo
4	1219.8	99.7	1654	10	AF128888	AF128888 Mus muscu
5	867.6	70.9	1498	10	AF178974	AF178974 Rattus no
6	672.8	55.0	1268	6	AR201144	AR201144 Sequence
7	653.4	53.4	1140	9	AF274571	AF274571 Homo sapi
8	641.8	52.4	1086	6	AR432047	AR432047 Sequence
9	641.8	52.4	1086	6	BD091715	BD091715 Novel deo
10	512.2	41.8	1029	9	AF333389	AF333389 Homo sapi
11	460.4	37.6	2337	10	AF33460386	AF33460386 Mus muscu
12	460.4	37.6	197457	2	AC127571	AC127571 Mus muscu
13	460.4	37.6	211003	2	AC141632	AC141632 Mus muscu
14	460.4	37.6	255925	2	AC114618	AC114618 Mus muscu
15	308.2	25.2	231698	2	AC118117	AC118117 Rattus no
16	308.2	25.2	306870	2	AC098557	AC098557 Rattus no
17	255.6	20.9	19164	9	AF334602	AF334602 Homo sapi
18	255.6	20.9	164529	2	AC027606	AC027606 Homo sapi
19	255.6	20.9	205034	9	AL359273	AL359273 Human DNA
20	202.6	16.6	642	10	AF33460385	AF33460385 Mus muscu
21	179	14.6	1374	10	AF33460382	AF33460382 Mus muscu
22	160.4	13.1	719	10	AF33460384	AF33460384 Mus muscu
23	151.4	12.4	1825	10	AF33460381	AF33460381 Mus muscu
24	114.8	9.4	1403	10	RATSP6T7	D50689 Rat DNA for
25	97.4	8.0	86816	2	AL136113	AL136113 Homo sapi
26	91.6	7.5	7470	14	FPO6408	AJ006408 Fowlpox v
27	91.6	7.5	288539	14	AF198100	AF198100 Fowlpox v
28	84	6.9	1391	9	AB074101	AB074101 Actus tri
29	82.4	6.7	1357	9	AB074093	AB074093 Homo sapi
30	82.4	6.7	1388	9	AB074098	AB074098 Papio ham
31	82.4	6.7	1390	9	AB074097	AB074097 Hylobates
32	82.4	6.7	1391	9	AB074099	AB074099 Macaca mu
33	82	6.7	1089	10	AF33460383	AF33460383 Mus muscu
34	80.8	6.6	1387	9	AB074094	AB074094 Pan trogl
35	80.8	6.6	1388	9	AB074096	AB074096 Pongo pyg
36	80.8	6.6	1390	9	AB074100	AB074100 Macaca fa
37	79.2	6.5	1388	9	AB074095	AB074095 Gorilla g
38	63	5.1	42331	3	CERG45A03	AC084623 Caenorhab
39	60.8	5.0	1083	9	BT007047	BT007047 Homo sapi
40	60.8	5.0	1083	12	BT007514	BT007514 Synthetic
41	60.8	5.0	1575	6	AR337903	AR337903 Sequence
42	60.8	5.0	1593	9	AB004574	AB004574 Homo sapi
43	60.8	5.0	1601	9	BC010419	BC010419 Homo sapi
44	60.8	5.0	1751	9	AF045937	AF045937 Homo sapi
45	60.8	5.0	1920	9	AF047016	AF047016 Homo sapi

ALIGNMENTS

RESULT 1	AR201143	Sequence 1	1224 bp	DNA	linear	PAT 20-APR-2002
LOCUS	AR201143	Sequence 1	from patent US 6358723.			
DEFINITION	AR201143					
ACCESSION	AR201143					
VERSION	AR201143.1	GI:20252031				
KEYWORDS		Unknown.				
SOURCE		Unknown.				
ORGANISM		Unknown.				
REFERENCE		Unclassified.				
AUTHORS		1 (bases 1 to 1224)				
TITLE		Eastman, A. Richard. and Krieser, R. Joe.				
JOURNAL		Deoxyribonuclease II beta. proteins and cDNAs				
FEATURES		Patent: US 6358723-A 1 19-MAR-2002;				
		Location/Qualifiers				



```
QY 422 GGAACAGACGAGGGGTTCTGGCTGATACACTCTGTTCCTCCAAAGTTTCCCCAGTTTCATG 481
Db 595 GGAACAGACGAGGGGTTCTGGCTGATACACTCTGTTCCTCCAAAGTTTCCCCAGTTTCATG 654
QY 482 GCTATGAGTACCAACCTCGGGGAGCGATATGACAAACCGCATCTGCATCATTTCG 541
Db 655 GCTATGAGTACCAACCTCGGGGAGCGATATGACAAACCGCATCTGCATCATTTCG 714
QY 542 GATACAGCCAGTTTGAGGAAATAGATTTTTCAGCTCTTGTGTTTACAAACCAACATCTACA 601
Db 715 GATACAGCCAGTTTGAGGAAATAGATTTTTCAGCTCTTGTGTTTACAAACCAACATCTACA 774
QY 602 GCTGCTTCATTCAGACACCTTTTCATCTGAAACCTTATCTACATGCCCCGGAGTGTGCCA 661
Db 775 GCTGCTTCATTCAGACACCTTTTCATCTGAAACCTTATCTACATGCCCCGGAGTGTGCCA 834
QY 662 ACTCCAGTTTCTTAAAGATCCCTGTCCGGTACCTCGCTGAACCTGACCTCAGCCCGGTC 721
Db 835 ACTCCAGTTTCTTAAAGATCCCTGTCCGGTACCTCGCTGAACCTGACCTCAGCCCGGTC 894
QY 722 TAAACTTCCTCCATTTTGCAAAATCAAGTTTTCATCTGATGACATCTTTTACAGGATGGA 781
Db 895 TAAACTTCCTCCATTTTGCAAAATCAAGTTTTCATCTGATGACATCTTTTACAGGATGGA 954
QY 782 TAGCTCAAAAGTTGAAGACACATTTGTTAGCACAACCTTGGCAGAAAAGAAACAGAGC 841
Db 955 TAGCTCAAAAGTTGAAGACACATTTGTTAGCACAACCTTGGCAGAAAAGAAACAGAGC 1014
QY 842 TTCCCTTCAAACTGTTCCCTGCTTACCATGCTCTCAACATCAAGTCCATTTGGGTTAACTT 901
Db 1015 TTCCCTTCAAACTGTTCCCTGCTTACCATGCTCTCAACATCAAGTCCATTTGGGTTAACTT 1074
QY 902 CCAAGTCTTACTTCAAGTTCTGCCAAGACCATTTCCAAATGGTGTGTTTCCATAAAGGGCT 961
Db 1075 CCAAGTCTTACTTCAAGTTCTGCCAAGACCATTTCCAAATGGTGTGTTTCCATAAAGGGCT 1134
QY 962 CGGAAATCGCTGGACCTGCTGAGACCTAAATCGAAGCCTACACCAAGCCTTAAAGAG 1021
Db 1135 CGGAAATCGCTGGACCTGCTGAGACCTAAATCGAAGCCTACACCAAGCCTTAAAGAG 1194
QY 1022 GTGGAGGATTCATCTGTACAAAGATCACTACATTTACAGGACATTTCTAATAATTATATC 1081
Db 1195 GTGGAGGATTCATCTGTACAAAGATCACTACATTTACAGGACATTTCTAATAATTATATC 1254
QY 1082 TCCGTTATGGGTTCTGTAAAGTAACTCGGTGAAAGCCACACCTCTGTCTTGAAGAAACA 1141
Db 1255 TCCGTTATGGGTTCTGTAAAGTAACTCGGTGAAAGCCACACCTCTGTCTTGAAGAAACA 1314
QY 1142 CTGGCACTGGAACATCTCGCTTGGATCTGTTCTCATTAATATCAAGGCTTCTGAGTGAG 1201
Db 1315 CTGGCACTGGAACATCTCGCTTGGATCTGTTCTCATTAATATCAAGGCTTCTGAGTGAG 1374
QY 1202 CACAAGCTAGGCTCCCAATAAAG 1224
Db 1375 CACAAGCTAGGCTCCCAATAAAG 1397
```

```
RESULT 3
LOCUS BD091714 1652 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel deoxyribonuclease, gene encoding thereof and use thereof.
ACCESSION BD091714
VERSION BD091714.1 GI:22637325
KEYWORDS WO 0112793-A/1.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1652)
Tanuma, S. and Shiohawa, D.
Novel deoxyribonuclease, gene encoding thereof and use thereof
Patent: WO 0112793-A 1 22-FEB-2001;
SEIICHI TANUMA, DAISUKE SHIOKAWA
```

```
COMMENT OS Mus musculus (mouse)
PN WO 0112793-A/1
PD 22-FEB-2001
PF 01-MAY-2000 WO 2000JP002893
PR 17-AUG-1999 JP 99P 230870
PT SEIICHI TANUMA, DAISUKE SHIOKAWA
PC C12N9/22, C12N15/12, C12N5/10, C12N7/01, A61K31/195, A61K48/00 CC
FH Key Location/Qualifiers
FT CDS (213)..(1274).
FEATURES
source 1..1652
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
ORIGIN
Query Match 99.7%; Score 1219.8; DB 6; Length 1652;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1221; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 CCAGTCCCTGCATGGAATGAGGCCACAGATAGAAATGACAGCAAAAGCCCTCTAAGAA 61
Db 175 CACAGTCCCTGCATGGAATGAGGCCACAGATAGAAATGACAGCAAAAGCCCTCTAAGAA 234
QY 62 CAGTCTCTTCTTGTCTCTTCTTGGCTCTCTGGGTCCTGGGACACAGAAATCTCAT 121
Db 235 CAGTCTCTTCTTGTCTCTTCTTGGCTCTCTGGGTCCTGGGACACAGAAATCTCAT 294
QY 122 GCAGAAATGAATGATGGTGAAGCTGTGGAATCTTATCTTTTATAGTTACCCAAAGGA 181
Db 295 GCAGAAATGAATGATGGTGAAGCTGTGGAATCTTATCTTTTATAGTTACCCAAAGGA 354
QY 182 CTAGCAGGCAAGTGAAGGCGGGCTGCAGTACTGTACTGGAATCCACAGCAAA 241
Db 355 CTAGCAGGCAAGTGAAGGCGGGCTGCAGTACTGTACTGGAATCCACAGCAAA 414
QY 242 CTTGAAACAGAGGCTCTTACCTGATTACAGACACAGAGTCTCTGGGAGGACCTTAC 301
Db 415 CTTGAAACAGAGGCTCTTACCTGATTACAGACACAGAGTCTCTGGGAGGACCTTAC 474
QY 302 AGCATCTGTATGACACATATATCCAGATGACACAGCCCTATCTAATATACAGATG 361
Db 475 AGCATCTGTATGACACATATATCCAGATGACACAGCCCTATCTAATATACAGATG 534
QY 362 GTGTCCCTGGATCTGTGAATTAACAGACAGATGACATGCCAAAGTCTGTGGTAT 421
Db 535 GTGTCCCTGGATCTGTGAATTAACAGACAGATGACATGCCAAAGTCTGTGGTAT 594
QY 422 GGAACAGAACGAGGGGTTCTGGCTGATACACTCTGTTCCCAAGTTTCCCGAGTTTCATG 481
Db 595 GGAACAGAACGAGGGGTTCTGGCTGATACACTCTGTTCCCAAGTTTCCCGAGTTTCATG 654
QY 482 GCTATGAGTACCAACCTCGGGGAGCGATATGACAAACCGGCATCTGCATCACTTCG 541
Db 655 GCTATGAGTACCAACCTCGGGGAGCGATATGACAAACCGGCATCTGCATCACTTCG 714
QY 542 GATACAGCCAGTTTGAGGAAATAGATTTTTCAGCTCTTGTCTTACAAACCAACATCTACA 601
Db 715 GATACAGCCAGTTTGAGGAAATAGATTTTTCAGCTCTTGTCTTACAAACCAACATCTACA 774
QY 602 GCTGCTTCATTCAGACACCTTTTCATCTGAAACCTTATCTACATGCCCCGGAGTGTGCCA 661
Db 775 GCTGCTTCATTCAGACACCTTTTCATCTGAAACCTTATCTACATGCCCCGGAGTGTGCCA 834
QY 662 ACTCCAGTTTCTTAAAGATCCCTGTCCGGTACCTCGCTGAACCTGACCTCAGCCCGGTC 721
Db 835 ACTCCAGTTTCTTAAAGATCCCTGTCCGGTACCTCGCTGAACCTGACCTCAGCCCGGTC 894
QY 722 TAAACTTCCTCCATTTTGCAAAATCAAGTTTTCATCTGATGACATCTTTTACAGGATGGA 781
Db 895 TAAACTTCCTCCATTTTGCAAAATCAAGTTTTCATCTGATGACATCTTTTACAGGATGGA 954
QY 782 TAGCTCAAAAGTTGAAGACACATTTGTTAGCACAACCTTGGCAGAAAAGAAACAGAGC 841
```





```

Db      1135  CCGCAATCGCTGGACCTGGCACTGGAGACCTAAATCGAAGCTACACCAAGCCTTAAGAG 1194
Qy      1022  GTGAGGATTCATCTGTACAAAGAAATCACTACATTTTACCAGGCATTTTCAATAAATATATC 1081
Db      1195  GTGAGGATTCATCTGTACAAAGAAATCACTACATTTTACCAGGCATTTTCAATAAATATATC 1254
Qy      1082  TCCGTTATGGTTCCTGTAAAGTAACTCGGTGAAGGCCACACCCCTCTGTCTCTGAAAAACA 1141
Db      1255  TCCGTTATGGTTCCTGTAAAGTAACTCGGTGAAGGCCACACCCCTCTGTCTCTGAAAAACA 1314
Qy      1142  CTGCACTGGAACATCTCGCTTGGATCTGTCTCCATAATATCAAGCTTCTGAGTGAG 1201
Db      1315  CTGCACTGGAACATCTCGCTTGGATCTGTCTCCATAATATTTCAAGCTTCTGAGTGAG 1374
Qy      1202  CACAACGTAGCTCCCAATAAAG 1224
Db      1375  CACAACGTAGCTCCCAATAAAG 1397

RESULT 5
AF178974
LOCUS      1498 bp mRNA linear ROD 29-NOV-1999
DEFINITION Rattus norvegicus deoxyribonuclease LIAD mRNA, complete cds.
ACCESSION AF178974
VERSION AF178974.1 GI:6470130
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
REFERENCE 1 (bases 1 to 1498)
AUTHORS Tanuma,S. and Shiohawa,D.
TITLE Cloning of a cDNA encoding a rat DNase II-like acid DNase
JOURNAL Biochem. Biophys. Res. Commun. 265 (2), 395-399 (1999)
MEDLINE 20025354
PUBMED 10558878
REFERENCE 2 (bases 1 to 1498)
AUTHORS Shiohawa,D. and Tanuma,S.
TITLE Direct Submission
JOURNAL Submitted (19-AUG-1999) Biochemistry, Science University of Tokyo,
Shinjuku-ku Ichigaya Funagawara-machi 12, Tokyo 162-0826, Japan
FEATURES
     source
         1..1498
             location/qualifiers
                 /organism="Rattus norvegicus"
                 /mol_type="mRNA"
                 /strain="Wistar"
                 /db_xref="taxon:10116"
                 159..1229
             /codon_start=1
             /product="deoxyribonuclease LIAD"
             /protein_id="AA113596.1"
             /db_xref="GI:6470131"
             /translation="MTAQLKAAFLFVLSGLVTPVISINEDGKAVDWFAFYKL
PARTSRGTMGLDLYLDSTNRWTSKSHLINSRSLRGTLEOLYRAHNAKNDTAY
LYNDAPVSNYNGYGHAKLLVNRVQGFLLHSIPKPPVPEKGVYESSROY
AQSGCLITKTSQYETIDSLVFPNIYSCFIPNIERWELHMPQMKASKLPS
RLTLVLSAOGNLFHFAKSTFYTDIDJPAWIAQKLKVHLLVESHQRKNHLPNCSL
PVHVNIIKLRGPLSDPFSHDHDKWCVSTKDSQARWTCIGDLNRSFQALRSGFTI
CSKNRIYIQSFDRLVSHYASCN"
CDS

```

```

Query Match      70.9%; Score 867.6; DB 10; Length 1498;
Best Local Similarity 82.8%; Pred. No. 8.6e-249;
Matches 1017; Conservative 0; Mismatches 204; Indels 7; Gaps 2;

Qy      2    CCGAGTCCCTGCATGAATGAAGGCCACAGATAGAAAATGACAGCAAGCCTCTAAAGAA 61
Db      121  CACAGTCTGAGCATGGAATGAAGGCCACAGACAGAAAATGACAGCCAGCCTCTAAAG 180
Qy      62  CAGTCTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 121
Db      181  CAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240

```

```

Qy      122  GCAGAAATGAATATATGTGAAGCTGTGACATGGTTTATCTTTTATAGTTACCCAAAAGGA 181
Db      241  GCATAAAGCAAGATGTGTAAGCCGTGACATGGTTTATCTTTTATAGTTACCCAGAAGGA 300
Qy      182  CTAGCAGGCAAGTGAAGAGCGCGGCTGCAGTACCTGTACCTGTGAGTCTCCACAGACAA 241
Db      301  CAGCAGAGAGGTACAGGGATGGGGCTGGATACCTGTACCTGTGAGTCTCCACATAGAA 360
Qy      242  CCTGGAACAAGAGCCTTACCTGATTAACAGCACAGGAGTGTCTCTGGGAGGACCTTTAC 301
Db      361  CCTGGACCAAGAGCCACCACCTGATTAACAGCAGCAGAGAGCTCCCTGGGAGGACCTTGG 420
Qy      302  AGCATCTGTATGACACACATTAATTCACGAATGACACAGCAGCCTATCTTAATATACAA 361
Db      421  AGCAGCTGTATGAAGCACACAATGCCAAGATGACACAGCCTATCTGATATACAATGACG 480
Qy      362  GTGTCCCTGGATCTGTGAATTTACAGCAGACAGATATGACATGCAATGCAAGGTCTGCT 421
Db      481  CTGTCCCTGGATCTGTGAATTTACAGCGGAATTTACGGGGCATGCCAAGGTCTGCTGGTAT 540
Qy      422  GGAACAGAACGCGAGGGGTTCTGGCTGATACACTCTGTTCCTCAAGTTTCCCCAGTTTCC 478
Db      541  GGAACAGAGTGCAGGGGTTCTGGCTGATTCATTCTATTTCCCAAGTTTCCCCAGTTTCCG 600
Qy      479  ---ATGCTATGATACCCAACTCCGGGAGGCGATATGGAACAAACCGGCATCTGCATCA 535
Db      601  AAAAGGCTATGATGAATACCCAAAGCTCGGGGAGGCATATGCAAAAGTGGGCTCTGCATCA 660
Qy      536  CTTTCCGATACAGCCAGTTTGAAGAAATAGATTTTCAGCTCTTGGTCTTTCACACCAACA 595
Db      661  CTCTCAATACAGCCAGTATGACAAATAGATTTCTCAGCTCTTGGTCTTTCACGCCAAACA 720
Qy      596  TCTACAGTCTGTTCAATTCCAAGCACCTTTCACTGGAACTTATCTACATGATGATGATGAT 655
Db      721  TCTACAGTCTGTTTCAATTCCTCAACATCTTTCGGTGGAACTCATCCACATGATGATGAT 780
Qy      656  GTGCCAACTCCAGTCTCTTAAAGATCCCTGTGCGGTACCTGCTGAACCTGCACTCAGCCC 715
Db      781  GTGCCAAGTCCAGTCTGCTCAAAAGATCCCTGAGCGGCGCTCACTGTACTTCACTCAGCCC 840
Qy      716  AGGCTCTAAACTTCGTCATTTTGAAATCAAGTTTATCTGATGATGATGATGATGATGAT 775
Db      841  AGGCTCTAAACTTCCTCCATTTTGAAATCAACTTTTATCTGATGATGATGATGATGATGAT 900
Qy      776  GATGATAGTCTCAAAAGTTGAAGACACATTTGTAGACAAACCTGGGCAAGAAACAAAC 835
Db      901  CCTGGATAGTCTCAAAAGCTGAAGGTGCTTTGCTGTAGATCTCTGGCAGCGAAGAAC 960
Qy      836  AAGAGCTTCTTCAAACTGTTCCCTGCTTACCATGTCTTACAAATCAAGTCCATTTGGGG 895
Db      961  ACGAGCTTCTTCAAACTGTTCCCTGCTTACCATGTCTTACAAATCAAGGCAATTCGGG 1020
Qy      896  TAACTTCAAGTCTTACTTCACTGCTCCGCAAGACCATTCGAAATGGTGTGTTCATTA 955
Db      1021  GACCTCTGAGTCCGACATCCCTTCTCATCAGACCATTCGAAATGGTGTGTGTTCACAA 1080
Qy      956  AGGGCTCCGCAAAATCGCTGGACCTGCAATTTGAGAGACTTAAATCGAAGCCTTACCAAG 1015
Db      1081  AGGACTCTCAGGCTCGCTGGACCTGTATTGAGACCTAATTCGAGAGCCACATCAAGCCT 1140
Qy      1016  TAAGAGTGGAGGATTCATCTGTACAAAGATCACTACATTTACAGGCATTTTCAATAAT 1075
Db      1141  TGAGAAGTGGAGGATTTATCTGTTTCAAGAAATCGGTACATTTTACAGTCAATTTGATAG 1200
Qy      1076  TATATCTCCGTTATGGGTTCTGTAGTAACTCGGTGAAGGCCACACCTCTGTCTCTG 1135
Db      1201  TAGTTTCCATATATGCTCTCTGTAACTGAGTCTGGGAAAGGTCCCACTGCTGCTTGA 1260
Qy      1136  AAAACACTGGCACTGGAAACATCTCGCCTTGGATCTGTTCTCATATATCAAGGCTCTG 1195
Db      1261  AAGACACTGGCTCTGGGACATCTTGCCTTGGATCTATTTCTCATATATGTTAAAGTTTCT 1320

```

QY 1196 AGTGACCAACCTAGCTCCATRAAA 1223  
 Db 1321 AGTG-TCCTAACATCTCATCCATAAAA 1347

RESULT 6  
 AR201144  
 LOCUS 1268 bp DNA linear PAT 20-APR-2002  
 DEFINITION Sequence 3 from patent US 6358723.  
 ACCESSION AR201144  
 VERSION AR201144.1 GI:20252032  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 UNCLASSIFIED.  
 REFERENCE 1 (bases 1 to 1268)  
 AUTHORS Eastman,A.Richard, and Krieser,R.Joe.  
 TITLE Deoxyribonuclease II.beta. proteins and cdnas  
 JOURNAL Patent: US 6358723-A 3 19-MAR-2002;  
 FEATURES Location/Qualifiers  
 1..1268  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN  
 Query Match 55.0%; Score 672.8; DB 6; Length 1268;  
 Best local similarity 75.0%; Pred. No. 2.6e-190;  
 Matches 871; Conservative 0; Mismatches 282; Indels 9; Gaps 2;

QY 31 AGATAGAAATGACAGCAAGCCCTAAGAACAGTCTTCTTGTGCTCTCTTGTGCGCTC 90  
 Db 37 AAACAGAAATGATGGCAAGACTCTAAGAACATCTTGTGCTCTCTTGTGCGCTC 96  
 QY 91 TCTGGGTCTCTGGGACACACCAAGAAATCTCAATGCAAGAAATGAATATGATGAGCTGTGGAC 150  
 Db 97 TTCTGGGTCTCTGGGACACCAATCTCAATGCAAGAAATGAAGGAAAGCTGTGGAC 156  
 QY 151 TGGTTATCTTTTATATAGTTACCAAGAGCTAGCAAGCAAGTGAAGCGCGGGCTG 210  
 Db 157 TGGTTATCTTTTATAGTTACCAAGAGCTAGCAAGCAAGTGAAGCGCTGAGCTGAGTTA 216  
 QY 211 CAGTACCTGTACCTGGAGCTCCACAGCAAAACCTCGGAACAGAGCTCTACCTGATTAAC 270  
 Db 217 GAGTACCTGTACCTAGACTCTACACTAGAGCTGGAGAGAGTGAACCACTAATGAAT 276  
 QY 271 AGCAACAGAGTGTCTGGGAGGACCTTACAGATCTGTATGACACACATAATTCACAG 330  
 Db 277 GACACCAAGAGTGTCTGGGAGGACATTAACAGAGCTATATGAAGCATATGCGCTCTAAG 336  
 QY 331 AATG---ACAGAGCTATCTAATATACAGATGCTGCTCGGATCTGTGAATTTACAGC 387  
 Db 337 AGTAACACACAGAGCTATCTAATATACAGATGCTGCTCGGATGCTTAAACCTGTGAATTTACAGT 396  
 QY 388 AGACAGTATGACATGCCAAGAGTCTGCTGTATGAAACAGAGCAAGCGGGTCTTGCTG 447  
 Db 397 AGAAGTATGACACACCAAGGTTTACTGTGTGAAACAGAGTTCGAAGGTTCTGGCTG 456  
 QY 448 ATACACTGTCTTCCAGTTTCCCGAGTTC-----ATGGGTATGAGTACCCACCTCG 501  
 Db 457 ATTCAATCCATCCCTCAGTTTCCCTCCAATTCGGGAAGAGGCTATGATTATCCACCCACA 516  
 QY 502 GGGAGGCATATGACAAACCGGATCTGCATCACTTTTCGGATACAGCCAGTTTCAGGAA 561  
 Db 517 GGGAGCAAAATGGCAAGTGGCAATCTGATCACTTTCAAGTACACACAGTATGAGCA 576  
 QY 562 ATAGATTTTTCAGTCTTGTGCTTTACACCAACCAATCTACAGCTGTTTCATTCACAGACC 621  
 Db 577 ATAGATTTTCAGTCTTGTGCTGCAACCCCAAGCTATAGTCTTCCATCCAGCCACC 636  
 QY 622 TTTCATGTGAATCTTATCTATGATGCCCGGATGTGTGCAACTCCAGTTCCTTAAAGATC 681  
 Db 637 TTTCACAGAGCTCATTCATGATGCCCGGAGTGTGCAACCGAGGGCCAGCTCATCAGAGATT 696

QY 682 CTTGTCGGTACCTCGCTGAACCTGCACTCAGCCAGGGTCTAAACTTGTGCAATTTTGA 741  
 Db 697 CTTGCGAGGCTCTCTACCACTCTTCACTCGGCGAGGACAAATAATCTCCATTTTGA 756  
 QY 742 AATCAAGTTTATATCTGATGACATCTTACAGGATGGATAGCTCAAAAGTTGAAGACA 801  
 Db 757 AAGTCGGATCTTTCTTGAAGACATCTTTGCGAGCTTGGATGGCTCAACGGCTGAGACA 816  
 QY 802 CATTTGTTAGCACAAACCTGGCAGAAAAAGAAAAAGAGAGCTTCTCTTCAAACTGTTCCTG 861  
 Db 817 CACTTGTATACAGAAACCTGGCAGCGAAGAAAGACAGAGCTTCTCTTCAAACTGTTCCTG 876  
 QY 862 CTTTACCATGTCTACCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 921  
 Db 877 CTTTACCATGTCTACCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 936  
 QY 922 CGCAAGACCATTTCCAAATGGTGTGTTTCCATAAGGGCTCCGCAATCGCTGACCTGC 981  
 Db 937 TATCAAGATCACCGCAAGTGGTGTATTTCCCAAGGGCAACCAAAATCGCTGGACATGT 996  
 QY 982 ATTGGAGACCTTAATCGAAGCTTACACCAAGCTTGAAGAGTGGAGGATTCATCTGTACA 1041  
 Db 997 ATTGGAGACCTTAATCGAAGCTTACACCAAGCTTGAAGAGTGGAGGATTCATCTGTACC 1056  
 QY 1042 AAGAATCATCATTTTACCAAGGATTTCAATAATTTATCTCGTTATGGGTTCTGTAAAG 1101  
 Db 1057 CAGAATTTGGCAATTTTACCAAGGATTTCAAGGATTTAGTATTATATGAAAGCTGTAAAG 1116  
 QY 1102 TAACTCGGTGAAGGACACACCTCTGCTTGTGAAACACTGGCACTGGAAACATCTCGC 1161  
 Db 1117 TAAACTTGGTGAAGGACACAGGATCTATCATTTGAAACCTTGACAAATGGGTTCTTCTCC 1176  
 QY 1162 CTTGGATCTGTCTTCCATAATA 1183  
 Db 1177 ATTACACCTTTTATATTATA 1198

RESULT 7  
 AF274571 1140 bp mRNA linear PRI 12-JUN-2001  
 LOCUS Homo sapiens deoxyribonuclease II beta mRNA, complete cds.  
 DEFINITION AF274571  
 ACCESSION AF274571.1 GI:8571463  
 VERSION AF274571.1  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1140)  
 AUTHORS Krieser,R.J., MacLea,K.S., Park,J.P. and Eastman,A.  
 TITLE The cloning, genomic structure, localization, and expression of human deoxyribonuclease IIbeta  
 JOURNAL Gene 269 (1-2), 205-216 (2001)  
 MEDLINE 21272509  
 PUBMED 11376952  
 REFERENCE 2 (bases 1 to 1140)  
 AUTHORS Krieser,R.J. and Eastman,A.  
 TITLE Direct Submission  
 JOURNAL Submitted (18-MAY-2000) Pharmacology, Dartmouth Medical School,  
 7650 Remsen, Hanover, NH 03755, USA  
 FEATURES Location/Qualifiers  
 1..1140  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="1"  
 /map="lp22.3"  
 /note="additional 3'-untranslated sequence up to a potential polyadenylation site is present in several EST clones in the database"  
 46..1119  
 /note="acid endonuclease"  
 /codon\_start=1

CDS

```

/product="deoxyribonuclease II beta"
/protein_id="AAF76893.1"
/db_xref="GI:8571464"
/translation="MMARLARTSFALLFLGVLGAATISCRNEKAVDWTFYKL
PKQKSGEFTGLLEYLYLDSTTSWRKSEQLMDTKSVLGRITQOLLYEAVASKNNTA
YLYNDVGFAPVNSRYKIGHTKGLLWNVQGWFLHISIPQFPPIPEEGDYDPTTGR
NGSGICITPKYQYEAIDSLVNPVYSCSIPATFHQELIHPOLCTRASSSEIP
GRLLTTLQSAQGGKFLHFAKSDSFLDDIFAAWMAQRLLKHLLETWOKRQELPSNCS
LPYHVNKIKALSRHSYFSSYODHAKWCIQKGTNRWICIGDLNRSFHOAFRSQGF
ICTQNMQIYQAFQGLVLYESCK"

```

ORIGIN

```

Query Match      53.4%; Score 653.4; DB 9; Length 1140;
Best Local Similarity 75.9%; Pred. No. 1.8e-184;
Matches 836; Conservative 0; Mismatches 256; Indels 9; Gaps 2;

QY 31 AGATAGAAAATGACAGAAAGCTCTAAGAACAGTCTCTTTCTTTCCTCTCTTTGCGCCTC 90
DB 37 AAACAGAAAATGATGGCAAGACTGCTAAGAACATCTCTTTGCTTCTCTCTTGGCCTC 96

QY 91 TCTGGGCTCTGGGGACACACAGAAATCTCATGAGAAATGAATATGATGAGCTGTGGAC 150
DB 97 TTTGGGCTGTGGGGCAGCAACAATTTCAAGAGAAATGAGAGGAAAGCTGTGGAC 156

QY 151 TGGTTATCTTTTATAGTTACCAAAAGGACTAGCAAGCAAGTGAAGAGGGGGGCTG 210
DB 157 TGGTTTACTTTTATAAGTTTACCTAAAGACAAAACAAGGAAAGTGGAGAGCTGGGTTA 216

QY 211 CAGTACTGTACTGAGCTCCACAGACAACTGGAACAGAGCCCTCTACCTGATTAAC 270
DB 217 GAGTACTGTACTGAGCTCTCAACTAGAAAGCTGGAGGAAAGTGAAGCAATTAATGAAT 276

QY 271 AGCACCAGGAGTCTCTGGGAGGACCTTACAGCATCTGTATCACACACATATTCACG 330
DB 277 GACACCAAGAGTGTTTGGAGGACATTTACACAGCTATATGAGCAATATGCTCTAAG 336

QY 331 ---AATGACACAGCTTATCTAATATACAGATGAGTGTCCCTGGATCTGTGAATTACAGC 387
DB 337 AGTAAACAACACAGCTTATCTAATATACATGATGAGTCCCTAAACCTGTGAATTACAGT 396

QY 388 AGACAGTANGGACATGCCAAGTCTGCTGGTATGAAACAGACGAGGGTCTGGCTG 447
DB 397 AGAAGTATGGACACACCAAGGTTTACTGCTGTGGAAACAGAGTTCAAGGGTTCTGGCTG 456

QY 448 ATACACTCTGTCTCCCAAGTTTCCCCAGTTC-----ATGGCTATGAGTACCCCACTCG 501
DB 457 ATTCAATTCATCCCTCAGTTTCTCAATTCGGAAGAGGCTATGATATCCACCCACA 516

QY 502 GGGAGCGGATATGGACAAACCGGCATCTGCATCACTTTCCGATACAGCCAGTTTGAAGAA 561
DB 517 GGGAGACGAAATGGACAAAGTGGCATCTGCATAACTTTCAAGTACAAACAGTATGAGCA 576

QY 562 ATAGATTTTTCAGCTCTGTGCTTTACACCAACAATCTACAGTCTGCTTATTCAGACACC 621
DB 577 ATAGATTTTTCAGCTCTGTGCTTTCAGCCCAACCGTCTATAGCTGCTTCCATCCAGCCAC 636

QY 622 TTTCACTGAAACTTATCTATCATGCTCCCGAGTGTGCAACTCCAGTTTCTTAAAGATC 681
DB 637 TTTCACTGAGGCTTATCTATCATGCTCCCGAGTGTGCAACCGGCGCAGCTCATCAGAT 696

QY 682 CTGTCCGGTACTCTGCTGAACTGCACTCAGCCCGAGGCTTAAACTCTGTCATTTTGA 741
DB 697 CTGGCAGGCTCTCTACCACTTTCAGTCCGCCCGGACAAAATTCCTCCATTTTGA 756

QY 742 AAATCAAGTTTTTATCTATGATGATCTTTACAGATGATGATAGCTCAAAAGTTGAAGACA 801
DB 757 AAGTCGGAATTTCTTTTGAACACATCTTTGACGCTGGATGGCTCAACCGCTGAAGACA 816

QY 802 CATTTGTTAGCAAAACCTGGCAGAAAAGAAACAGAGCTTCTTCAAACTGTTCCCTG 861
DB 817 CACTTGTAAACAGAAACCTGGCAGGAAAGACAGAGCTTCTTCAAACTGTTCCCTG 876

QY 862 CTTTACCAGTCTTACACATCAAGTCCATTTGGGGTAACTTCCAAAGTCTTACTTCACTTCT 921

```

```

DB 877 CCTTACCATGTCTACATATATAAAGCAATTAATATACAGACACTCTTATTTCACTTCT 936
QY 922 CGCAAGACCAATTCCTCAATGCTGTGTTTCCATAAAGGGCTCCGCAAAATCGCTGGACCTGC 981
DB 937 TATCAAGATCAGCGCAAGTGTGTATTTCCAAAAGGGCACCAAAATCGCTGGACATGT 996
QY 982 ATTGGAGACCTAATCGAAGCCTTACCAAGCCTTAAGAGGTGAGGATTCATCTGTACA 1041
DB 997 ATTGGAGACCTAATCGAAGCCTTCCACCAAGCCTTCCAGAAAGTGGAGGATTCATTTGTACC 1056

QY 1042 AAGAATCACTACATTTTACCAGGCAATTTATAAATATATATATCCGTTATGGTCTGTAAAG 1101
DB 1057 CAGATTTGGCAATTTTACCAGGATTTTCAAGGATTTAGTATATATACTAATAAGCTGTAAAG 1116

QY 1102 TAAACTCGGTGAAGGCCACA 1122
DB 1117 TAAACTTGGTGAAGGACACA 1137

RESULT 8
AR432047
LOCUS
DEFINITION
Sequence 4 from patent US 6653118.
AR432047
ACCESSION
AR432047.1
VERSION
GI:40194250
KEYWORDS
SOURCE
Unknown.
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 1086)
Tanuma,S.-i. and Shiohawa,D.
Deoxyribonuclease, Gene encoding same and use thereof
Patent: US 6653118-A 4 25-NOV-2003;
JOURNAL
Location/Qualifiers
FEATURES
1..1086
source
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Query Match      52.4%; Score 641.8; DB 6; Length 1086;
Best Local Similarity 75.9%; Pred. No. 5.4e-181;
Matches 822; Conservative 0; Mismatches 252; Indels 9; Gaps 2;

QY 31 AGATAGAAAATGACAGAAAGCTCTAAGAACAGTCTTTTCTTCTTCTTCTTCTTCTTCTTCT 90
DB 4 AAACAGAAAATGATGGCAAGACTGTGTAAGAACATCTTTTGTCTTCTTCTTCTTCTTCTTCT 63

QY 91 TCTGGGCTCTGGGGACACCAAGAAATCTCATGAGAAATGAATATGATGAGCTGTGGAC 150
DB 64 TTTGGGCTCTGGGGCAGCAACAATTTTCATGAGAAATGAAAGAGGAAAGCTGTGGAC 123

QY 151 TGGTTTATCTTTTATAAGTTTACCCAAAAGGACTAGCAAGGCAAGTGAAGAGCGGGGCTG 210
DB 124 TGGTTTACTTTTATAAGTTTACTTAAAGACAAAACAGAGAAAGTGGAGAGACTGGGTTA 183

QY 211 CAGTACTGTACTGGACTTCCAAAGACAAACCTCGAACAAGAGCTCTTACCTGATTAAC 270
DB 184 GAGTACTGTACTGAGACTCTACAACTAGAAGCTGGAGGAGAGTGAAGCAACTAATGAAT 243

QY 271 AGCACCAGGAGTCTCTGGGAGGAGCTTACAGCATCTGTATGACACACATTAATTTCCAGC 330
DB 244 GACACCAAGAGTGTTTTGGGAAGGACATTAACAACAGCTATATGAAGCATATGCTCTAAG 303

QY 331 ---AATGACACAGCTTATCTAATATACAAACGATGTGTCCCTGGATCTGTGAATTACAGC 387
DB 304 AGTAAACAACAGCTTATCTAATATACAAATGAGGAGTCCCTTAAACCTGTGTAATTACAGC 363

QY 388 AGACAGTATGACATGCAAAAGTCTGCTGGTATGGAACAGAAACGAGGGTCTTGGGCTG 447
DB 364 AGAAAGTATGACACACCAAGGTTTACTGCTGTGGAACAGAGTTTCAAGGGTCTTGGCTG 423

QY 448 ATACACTCTGTCTCCCAAGTTTCCCCAGTTC-----ATGGCTATGAGTACCCAACTCG 501

```

Db	424	ATTCAATCCATCCCTCAGTTCTCCATTCCTCGAAGAGGCTATGATATCAACCACA	483
Qy	502	GGGAGGCGATATGACAAACCGCATCTGCATCACTTCGGATACAGCCAGTTGGAGAA	561
Db	484	GGGAGGCGAATGGACAAAGTGGCATCTGATTAATCTTCAAGTACACAGATATGAGCA	543
Qy	562	ATGATTTTCAGCTCTTGGTCTTACACCAACATCTACAGCTGCTTCATTCACAGACC	621
Db	544	ATGATTTTCAGCTCTTGGTCTTACACCAACATCTTATAGCTGCTCCATCCACAGACC	603
Qy	622	TTTCACTGGAACTTATCTACATCCCGGATGTCGCAACTCCAGTTCCTTAAAGATC	681
Db	604	TTTCACTGGAACTTATCTACATCCCGGATGTCGCAACTCCAGTTCCTTAAAGATC	663
Qy	682	CTGTTCGGTACCTCGTGAATGCACTCAGCCAGGGTCTAAATCTTCGTCATTTTGA	741
Db	664	CTGTTCGGTACCTCGTGAATGCACTCAGCCAGGGTCTAAATCTTCGTCATTTTGA	723
Qy	742	AAATCAAGTTTATATCTGATGATCTTTACAGATGATAGCTCAAAAGTTGAAGCA	801
Db	724	AAATCAAGTTTATATCTGATGATCTTTACAGATGATAGCTCAAAAGTTGAAGCA	783
Qy	802	CAATTTGTTAGCAGAACTTGTGATGATGTCGCAACTCCAGTTCCTTAAAGATC	861
Db	784	CAATTTGTTAGCAGAACTTGTGATGATGTCGCAACTCCAGTTCCTTAAAGATC	843
Qy	862	CTTACCATGCTCAACATCAAGTCCATTCAGGCTTCCCTTCAAACTGCTCCCTT	921
Db	844	CTTACCATGCTCAACATCAAGTCCATTCAGGCTTCCCTTCAAACTGCTCCCTT	903
Qy	922	CGCAGACCATTTCCAAATGGTGTGTTTCCATAAAGGGCTCCGCAAAATCGTGACCTGC	981
Db	904	TATCAAGATCATGCGCAAGTGTGTATTTCCCAAAAGGGCACTTCAAACTGCTCCCTT	963
Qy	982	ATTGAGACCTTAAATCGAAGCTTACCAAGCTTAAAGAGTGGAGGATTCATCTGACA	1041
Db	964	ATTGAGACCTTAAATCGAAGCTTACCAAGCTTAAAGAGTGGAGGATTCATCTGACA	1023
Qy	1042	AAGATCATCTACATTTACAGGATTTTCAATTAATATCTCCGTTATGGTCTGTAAG	1101
Db	1024	CAGAAATGGCAATTTACAGGATTTTCAAGGATTTAGTATTTATATGAAAGCTGTAAG	1083
Qy	1102	TAA 1104	
Db	1084	TAA 1086	
RESULT 9			
BD091715		1086 bp DNA linear	PAT 27-AUG-2002
LOCUS			
DEFINITION		Novel deoxyribonuclease, gene encoding thereof and use thereof.	
ACCESSION		BD091715	
VERSION		BD091715.1	GI:22637326
KEYWORDS		WO 0112793-A/2.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE			
AUTHORS		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
TITLE		Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.	
JOURNAL		1. (bases 1 to 1086)	
COMMENT		Tanuma, S. and Shikawa, D.	
		Novel deoxyribonuclease, gene encoding thereof and use thereof	
		Patent: WO 0112793-A 2 22-FEB-2001;	
		SEIICHI TANUMA, DAISUKE SHIKAWA	
		OS Homo sapiens (human)	
		PN WO 0112793-A/2	
		PD 22-FEB-2001	
		PF 01-MAY-2000 WO 2000JP002893	
		PR 17-AUG-1999 JP 99P 230870	
		PI SEIICHI TANUMA, DAISUKE SHIKAWA	
		PC C12N15/22, C12N15/12, C12N5/10, C12N7/01, A61K31/195, A61K48/00 CC	
		FT Key	
		Location/Qualifiers	
		(1) . . (1083).	

FEATURES	Location/Qualifiers	
source	1. .1086	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
ORIGIN		
	Query Match 52.4%; Score 641.8; DB 6; Length 1086;	
	Best Local Similarity 75.9%; Pred. No. 5.4e-181;	
	Matches 822; Conservative 0; Mismatches 252; Indels 9; Gaps 2;	
Qy	31	AGATAGAAATGACAGCAAAAGCCTTAAGAACAGTCTTTCTTTGCTCTTTTGGCCCTC 90
Db	4	AAACAGAAATGATGGCAGACTGCTTAAGAACATCTTTGCTTTGCTCTTTTGGCCCTC 63
Qy	91	TCCTGGGTCTGGGGACACAGAAATCTCATGACAGAAATGAATATGATGTAAGCTGTGGAC 150
Db	64	TTTGGGTGCTGGGGGAGCAACAATTTCAITGACAAATGAAGAGGAAAGCTGTGGAC 123
Qy	151	TGTTTTATCTTTTATAGTTTACCCAAAGGACTAGCAAGCAAGTGAAGAGGGGGGCTG 210
Db	124	TGGTTTACTTTTATAGTTTACCTTAAAGACAAAACAAGGAAAGTGGAGAGACTGGGTTA 183
Qy	211	CAGTACCTGTACCTGGGACTCAGACAAACCTGGACACAGAGCCTCTACCTGATTAAC 270
Db	184	GAGTACCTGTACCTGAGACTCTACAACTAGAAGCTGGAGGAGAGTGAAGCACTAATGAAT 243
Qy	271	AGCACCAGGAGTGTCTGTTGGGAGGACCTTACAGCATCTGTATGACACACATTAATCCACG 330
Db	244	GACACCAGGAGTGTGTTGGGAGGACATTAACACAGCTATATGAAGCATATGCTCTAAG 303
Qy	331	---AATGACACAGCCTATTAATATACAGATGGTGTCTCTGGATCTGTGAATTAACG 387
Db	304	AGTAAACACACAGCCTATTAATATACATGATGAGTCCCTTAAACCTGTGAATTAACG 363
Qy	388	AGACAGTATGACATGCAAAAGCTCTGTGTATGGAACAGAAACGAGGGGTCTGGCTG 447
Db	364	AGAAAGTATGACACACCAAGGTTTACTGCTGTGGAACAGAGTTCAAGGGTCTGGCTG 423
Qy	448	ATACACTGTGTTCCTCAAGTTTCCCCAGTTTC-----ATGGCTATGAGTACCCACCTCG 501
Db	424	ATTCACTCCATCTCTCAGTTTCTCCATTCGGAAGAGGCTATGATTAATCCACCAACA 483
Qy	502	GGGAGGCGATATGACAAACCGGATCTGATATCACTTTTCGGATACAGCAGCTTTGAGGAA 561
Db	484	GGGAGGCGAATGGACAAAGTGGCATCTGCATAAATTTCAAGTACACACAGTATGAGCA 543
Qy	562	ATAGATTTTACGCTCTTGGTCTTACACCAACATCTACAGCTGCTTTCATTCACAGCACC 621
Db	544	ATAGATTTTACGCTCTTGGTCTTACACCAACATCTTATAGCTGCTTCCATCCAGCACC 603
Qy	622	TTTCACTGGAACTTATCTACATGCGCGGATGTCGCAACTCCAGTTCCTTAAAGATC 681
Db	604	TTTCACTGGAGCTCAITTCATGCGGCTGTGTCACAGGGCCAGCTCATCAGAGATT 663
Qy	682	CCTGTCCGGTACCTCGTGAATGCACTCAGCCAGGGTCTAAATCTTCGTCATTTTGA 741
Db	664	CCTGTCCGGTACCTCGTGAATGCACTCAGTCCGGCCCGGAGGACAAAATTCCTCCATTTGCA 723
Qy	742	AAATCAAGTTTATCTGATGACATCTTACAGGATGATAGCTCAAAAGTTGAAGCA 801
Db	724	AAGTCGGATTTCTTTTCTGATGGCACTTTGAGCTGGATGCTTCAACGGCTGAAGCA 783
Qy	802	CAATTTGTTAGCAGAACTTGGCAGAAAAAGAAACAGAGCTTCTTCAAACTGTTCCCTG 861
Db	784	CATTTGTTTACAGAAACCTGGCAGCGAAAAAGACAAAGAGCTTCTTCAAACTGCTCCCTT 843
Qy	862	CCTTACCATGCTCAACATCAAGTCCATTTGGGGTAACTTCAAGCTCTTCTTCAAGTCT 921
Db	844	CCTTACCATGCTCAACATTAAGCAATTAATTAATCAAGCACTCTTATTTTCAAGTCT 903
Qy	922	CGCCAGACCATTTCCAAATGGTGTGTTTCCATAAAGGGCTCCGCAAAATCGTGACCTGC 981

```

Db      904  TATCAAGATCATGCAAGTGGTGTATTTCCAAAAGGCGCACAAAATCGTGGACATG 963
QY      982  ATTGGAGACCTAAATCGAAGCTTACACCAAGCCTTAAGAGGTGGAGATTCATCTGTACA 1041
Db      964  ATTGGAGACCTAAATCGAGTCCACCAAGCCTTCAGAAGTGGAGATTCATTTGTACC 1023
QY      1042  AAGAACTACATACATTTACCAGGCAATTCVAAATATATATCTCCGTTATGGTCTCTGAAG 1101
Db      1024  CAGAAATGGCAAAATTTACCAAGCATTTCAAGGATAGTATATATATACATGAAAGCTGTAA 1083
QY      1102  TAA 1104
Db      1084  TAA 1086

RESULT 10
AF333389
LOCUS      Homo sapiens lung-specific DLAD mRNA, complete cds. PRI 22-NOV-2001
DEFINITION
ACCESSION AF333389
VERSION    AF333389.1 GI:17046270
KEYWORDS   .
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 1029)
AUTHORS   Shiohawa,D. and Tanuma,S.-I.
TITLE     Isolation and characterization of the DLAD/Dlad genes, which lie
          head-to-head with the genes for urate oxidase
JOURNAL   Biochem. Biophys. Res. Commun. 288 (5), 1119-1128 (2001).
MEDLINE   21556924
PUBMED    11700027
REFERENCE  2 (bases 1 to 1029)
AUTHORS   Shiohawa,D. and Tanuma,S.-I.
TITLE     Direct Submission
JOURNAL   Submitted (04-JAN-2001) Biochemistry, Science Univ. of Tokyo,
          Ichigaya, Shinjuku-Ku, Tokyo 162-0826, Japan
FEATURES   Location/Qualifiers
            1..1029
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             408..869
              /note="expressed only in lung; does not include first two
              exons"
             /cddon_start=1
             /product="lung-specific DLAD"
             /protein_id="AAL34448.1"
             /db_xref="GI:17046271"
             /translation="MPQLCTRASSSEIPGRLLTLTQSAOGKFLHFAKSDSFLDDIEA
             AWAQRLKTHLLTETWKRORLPNSCLPYHVYNIKALKLSHYSYFSYQDHAKWCI
             SQKGTKNRWTCIGDLNRSHPHQAFRSGGFICTQNWQIYQAPQGLVLYESCK"

ORIGIN
Query Match      41.8%; Score 512.2; DB 9; Length 1029;
Best Local Similarity 75.7%; Pred. No. 4.4e-142;
Matches 650; Conservative 0; Mismatches 203; Indels 6; Gaps 1;

QY      331  AATGACACAGCCTATCTAATATACACAGTGTGTCCCTGGATCTGTGAATTACAGCAGA 390
Db      90  AACACACACAGCCTATCTAATATACATGATGAGTCCCTTAAACCTGTGAATTACAGCAGA 149
QY      391  CAGTATGACATGCCAAAGTCTGTGGTATGGACAGACCGAGGCTCTGGCTGATA 450
Db      150  AAGTATGGACACACCAAGAGTTACTGTGTGGACAGAGTTCAAGGGTCTGGCTGATT 209
QY      451  CACTCTGTTCCCAAGTTTCCCCAGTTC-----ATGGCTATGATGATCCCAACCTCGGG 504
Db      210  CATTCCATCCCTCAGTTTCCTCCATTCGGAGAGAGCTATGATATCCACCCACAGGG 269
QY      505  AGGCGATATGGACAAACCGGCAATCTGCATCTTTCCGATACAGCCAGTTTGAGGAATA 564

```

```

Db      270  AGACGAATGACAAAGTGGCATCTGCATACATTTCAAGTACACACAGTATGAGGCAATA 329
QY      565  GATTTTCAGCTCTTGGTCTTACCAACCAACATCTACAGCTGCTTCAATCCAAAGCACCTTT 624
Db      330  GATTTCTCAGCTCTTGGTCTGCAACCCCAACGCTCTATAGCTGCTCCATCCAGCCACCTTT 389
QY      625  CACTGGAAACTTATCTACATGCCCGGATGTGCGCAACTCCAGTTCCTTAAAGATCCCT 684
Db      390  CACAGAGCTCANTCACATGCCCCAGCTGTGACCCAGGGCCAGCTCATCAGAGATTCT 449
QY      685  GTCGGTACCTCGCTGAACTGCACCTCAGCCAGGCTTAAACTTCGTCATTTTGCAAAA 744
Db      450  GCGAGGCTCTCACCACACTTCAGTCGCCCGCAGGACAAAATTCCTCCATTTTGCAAG 509
QY      745  TCAAGTTTTTATCTGATGACATCTTTACAGGATGATAGCTCAAAAGTTGAAGACACAT 804
Db      510  TCGGATTTCTTTCTTGACGACATCTTTGACGCTGGATGGCTCAACGGCTGAAGACACAC 569
QY      805  TTCTTAGCACAAACCTGGCAGAAAAGAAACAGAGCTTCCTTCAAACCTGTTCCCTGCT 864
Db      570  TTGTTAACAGAAACCTGGCAGCGCAAGAGCTTCCTTCAAACCTGCTCCCTTCCT 629
QY      865  TACCATGTCTACAAACATCAAGTCCATTTGGGGTAACTTCCAAAGCTTACTTCAGTTCTCGC 924
Db      630  TACCATGTCTACAAATATAAAGCAATTAATATCAGACACTTTATTTCACTTTCTAT 689
QY      925  CAAAGCATTCCAAATGGTGTGTTCCATAAAGGGTCCGCAAAATCGCTGGACCTGCATT 984
Db      690  CAAAGTATCATGCCAAGTGGTGTATTTCCCAAAGGGCACCAAAATCGCTGGACATGTA 749
QY      985  GGAGACCTAAATCGAGCCTCACCAAGCCTTAAGAGTGGAGGATTCATCTGTACAAAG 1044
Db      750  GGAGACCTAAATCGAGGCTCCACCAAGCCTTCAGAAGTGGAGGATTCATTTGTACCCAG 809
QY      1045  AATCACTACATTTACCAGGCATTTCAATAATATATCTCGTTATCGGTTCTCTAAGTAA 1104
Db      810  AATTCGCAATTTACCAAGCATTTCAAGGATAGTATATATCTAAGAAAGCTGTAAAGTAA 869
QY      1105  ACTCGTGAAGGCGCACACCCTCTGTCCTTGAACACATCGGCACCTGGAAACATCTCGCCT 1164
Db      870  ACTTGTGAAGGACACAGGTACTATCATTTGAAAACCTTGAACAATGGGTCTTCTTCCATT 929
QY      1165  GGATCTGTTCTCCATAATA 1183
Db      930  ACACCTTCTTTATATTTTA 948

RESULT 11
AF334603S6
LOCUS      Mus musculus deoxyribonuclease DLAD (Dlad) gene, exon 6 and
DEFINITION complete cds.
ACCESSION AF334608
VERSION    AF334608.1 GI:17046280
KEYWORDS   .
SEGMENT    6 of 6
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
REFERENCE  1 (bases 1 to 2337)
AUTHORS   Shiohawa,D. and Tanuma,S.-I.
TITLE     Isolation and characterization of the DLAD/Dlad genes, which lie
          head-to-head with the genes for urate oxidase
JOURNAL   Biochem. Biophys. Res. Commun. 288 (5), 1119-1128 (2001)
MEDLINE   21556924
PUBMED    11700027
REFERENCE  2 (bases 1 to 2337)
AUTHORS   Shiohawa,D. and Tanuma,S.-I.
TITLE     Direct Submission
JOURNAL   Submitted (09-JAN-2001) Department of Biochemistry, Science
          University of Tokyo, Ichigaya, Shinjuku-ku, Tokyo 162-0826, Japan
FEATURES   Location/Qualifiers

```

```
source      1. .2337
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /strain="C57BL/6"
            /db_xref="taxon:10090"
gene        order(AF334603.1:<1476. 1825,AF334604.1:1. 1374,
            AF334605.1:1. 1089,AF334606.1:1. 719,AF334607.1:1. 642,
            1. 1182)
            /gene="Dlad"
mRNA        join(AF334603.1:<1476. 1588,AF334604.1:665. 842,
            AF334605.1:630. 708,AF334606.1:484. 639,
            AF334607.1:102. 299,486. 1182)
            /gene="Dlad"
            /product="deoxyribonuclease Dlad"
CDS         join(AF334603.1:1476. 1588,AF334604.1:665. 842,
            AF334605.1:630. 708,AF334606.1:484. 639,
            AF334607.1:102. 299,486. 826)
            /gene="Dlad"
            /codon_start=1
            /product="deoxyribonuclease Dlad"
            /protein_id="AAL34450.1"
            /db_xref="GI:17046281"
            /translation="MTAKPLRTVLISLFFALSGVLGTPPEISCRNEYGEAVDFWFIYK
            PKTSKASBAGLYLDSTROTWNKSLYLINSALGRTQLQHLDYTHNSINDTAY
            LIYNDGVPVSRYOYGHAKGLVWNRTOGLIHSVPKPPVGHVYEVTSGRYGO
            TGICITPGYQPEIDRQLVLQPNLYSCFIPSTFHWKLIYMPRMCASSSLKIPVY
            LAELHSQGLNFVHFASKSFYDDIFFTGWIQAOKLKHLLAQTKWKKQKQELPNSGLPY
            HVNISKIGVTSKYSFSSRQDSKWCVSIRGSANRWTCIGLNRKSLHQALRGGFICT
            KNYHYQAFHKLXLYRGCK"
            /number=6
            /gene="Dlad"
            /number=6
exon
ORIGIN
Query Match      37.6%; Score 460.4; DB 10; Length 2337;
Best Local Similarity 99.8%; Pred. No. 1.6e-126;
Matches 461; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 763 GACATCTTTACAGATGATAGCTCAAAGTTGAAGACACATTTGTTAGCAACAACCTGG 822
Db 485 GACATCTTTACAGATGATAGCTCAAAGTTGAAGACACATTTGTTAGCAACAACCTGG 544
QY 823 CAGAAAAGAACAGAGCTTCCTTCAAACCTGTTCCCTGCTTACCAGTCTCAACAATC 882
Db 545 CAGAAAAGAACAGAGCTTCCTTCAAACCTGTTCCCTGCTTACCAGTCTCAACAATC 604
QY 883 AAGTCCATTGGGGTAACTTCAAGTCTTACTTCACTTCTGCTCCCAAGACCATTCCAAATGG 942
Db 605 AAGTCCATTGGGGTAACTTCAAGTCTTACTTCACTTCTGCTCCCAAGACCATTCCAAATGG 664
QY 943 TGTGTTTCCATAAAGGGTCCGCAATCGCTGGACCTGCTGAGACCTAAATCGAAGC 1002
Db 665 TGTGTTTCCATAAAGGGTCCGCAATCGCTGGACCTGCTGAGACCTAAATCGAAGC 724
QY 1003 CTACACCAAGCTTAAAGAGTGGAGGATTCATCTCTACAAAAGATCACTACATTTACCAG 1062
Db 725 CTACACCAAGCTTAAAGAGTGGAGGATTCATCTCTACAAAAGATCACTACATTTACCAG 784
QY 1063 GCATTTCAATAATATATCTCCGTTATCGGTTCTGTAAAGTAACTCGGTGAAGGCCACA 1122
Db 785 GCATTTCAATAATATATCTCCGTTATCGGTTCTGTAAAGTAACTCGGTGAAGGCCACA 844
QY 1123 CCCTCTGCTTGAACCACTGGCACTGGACATCTCGCTTGGATCTGTTCTCCATAAT 1182
Db 845 CCCTCTGCTTGAACCACTGGCACTGGACATCTCGCTTGGATCTGTTCTCCATAAT 904
QY 1183 ATCAAGGCTTCTGAGTGAGCACCACTAGCGTCCCAATAAAG 1224
Db 905 TTCAAGGCTTCTGAGTGAGCACCACTAGCGTCCCAATAAAG 946
RESULT 12
AC127571
```

```
LOCUS      AC127571
DEFINITION Mus musculus chromosome UNK clone RP24-227D4, WORKING DRAFT
SEQUENCE, 4 unordered pieces.
AC127571
VERSION    GI:24080752
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE     Mus musculus (house mouse)
ORGANISM   Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 197457)
            McPherson,J.D. and Waterston,R.H.
            The sequence of Mus musculus clone
            Unpublished
REFERENCE  2 (bases 1 to 197457)
            McPherson,J.D. and Waterston,R.H.
            Direct Submission
            Submitted (17-JUL-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE  3 (bases 1 to 197457)
            McPherson,J.D. and Waterston,R.H.
            Direct Submission
            Submitted (17-OCT-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
COMMENT    On Oct 17, 2002 this sequence version replaced gi:21887002.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: M_B80227D04
----- Summary Statistics -----
Sequencing vector: M13; 0%
Chemistry: Dye-primer ET; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 195668 bases at least Q40
Consensus quality: 196117 bases at least Q30
Consensus quality: 196344 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 31865: contig of 31865 bp in length
* 31866 31965: gap of unknown length
* 31966 70958: contig of 38993 bp in length
* 70959 71058: gap of unknown length
* 71059 136518: contig of 65460 bp in length
* 136519 136619: gap of unknown length
* 136619 197457: contig of 60839 bp in length.
            location/Qualifiers
FEATURES   1..197457
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /chromosome="UNK"
            /clones="RP24-227D4"
            misc_feature
            1..31865
            /note="assembly_name:Contig14"
            misc_feature
            31966..70958
            /note="assembly_name:Contig15"
            misc_feature
            71059..136518
            /note="assembly_name:Contig16"
            misc_feature
            136619..197457
```

```

ORIGIN
Query Match 37.6%; Score 460.4; DB 2; Length 197457;
Best Local Similarity 99.8%; Pred. No. 1.7e-126;
Matches 461; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

/Note="assembly_name:Contig17"

763 GACATCTTTACAGATGATAGTCTCAAAAGTCTGAAGACACATTTGTTAGCACAAACCTGG 822
Db GACATCTTTACAGATGATAGTCTCAAAAGTCTGAAGACACATTTGTTAGCACAAACCTGG 68166
QY 823 CAGAAAAGAAACAGAGCTTCTTCAAACTGTTCCCTGCTTACCAGTCTCAACATC 882
Db CAGAAAAGAAACAGAGCTTCTTCAAACTGTTCCCTGCTTACCAGTCTCAACATC 68226
QY 883 AAGTCCATTGGGGTAACTTCCAGTCTTACTCAGTCTCGCCAGACCATTTCAATGG 942
Db AAGTCCATTGGGGTAACTTCCAGTCTTACTCAGTCTCGCCAGACCATTTCAATGG 68286
QY 943 TGTGTTTCCATTAAGGGCTCCGCAATCGCTGACCTGCAATTCGAGACCTTAAATCGAAGC 1002
Db TGTGTTTCCATTAAGGGCTCCGCAATCGCTGACCTGCAATTCGAGACCTTAAATCGAAGC 68346
QY 1003 CTACACACAGCCTTAAGAGTGGAGATTCATCTGTACAAAAGATCACTACATTTACCAG 1062
Db CTACACACAGCCTTAAGAGTGGAGATTCATCTGTACAAAAGATCACTACATTTACCAG 68406
QY 1063 GCATTTCAATAATATATCTCCGTTATGGTCTGTAACTGCTGAAAGGCCACA 1122
Db GCATTTCAATAATATATCTCCGTTATGGTCTGTAACTGCTGAAAGGCCACA 68466
QY 1123 CCTCTGCTCTTGAACACTGSCACTGGAACATCTCGCTTGGATCTGTCTCCATAAT 1182
Db CCTCTGCTCTTGAACACTGSCACTGGAACATCTCGCTTGGATCTGTCTCCATAAT 68526
QY 1183 ATCAAGGCTTCTGAGTGAGACACAACTAGCTGCTCAATAAAG 1224
Db ATCAAGGCTTCTGAGTGAGACACAACTAGCTGCTCAATAAAG 68568

RESULT 13
AC141632/c 211003 bp DNA linear HTG 19-MAR-2003
LOCUS Mus musculus chromosome UNK clone RP24-100G24, WORKING DRAFT
DEFINITION SEQUENCE, 10 unordered pieces.
ACCESSION AC141632
VERSION AC141632.2 GI:29124286
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 211003)
McPherson,J.D. and Waterston,R.H.
The sequence of Mus musculus clone
Unpublished
2 (bases 1 to 211003)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (18-MAR-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 211003)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (19-MAR-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
On Mar 19, 2003 this sequence version replaced gi:29029422.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu

----- Project Information -----
Center project name: M_BB0100G24
----- Summary Statistics -----
Sequencing vector: M13; %
Chemistry: Dye-terminator Big Dye; % of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 208115 bases at least Q40
Consensus quality: 208752 bases at least Q30
Consensus quality: 209170 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

1 1455: contig of 1455 bp in length
* 1456: gap of unknown length
* 1556: contig of 5011 bp in length
* 6567: gap of unknown length
* 6567 10113: contig of 3447 bp in length
* 10114 10213: gap of unknown length
* 10214 19689: contig of 9476 bp in length
* 19690 31278: contig of 11489 bp in length
* 31279 31378: gap of unknown length
* 31379 51289: contig of 19911 bp in length
* 51290 51389: gap of unknown length
* 51390 81980: contig of 30491 bp in length
* 81981 124749: contig of 42769 bp in length
* 124750 124849: gap of unknown length
* 124850 164718: contig of 39869 bp in length
* 164719 211003: contig of 46185 bp in length.

FEATURES
Location/Qualifiers
1..211003
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP24-100G24"
1..1455
/note="assembly_name:Contig31"
1556..6566
/note="assembly_name:Contig32"
6567..10113
/note="assembly_name:Contig33"
10214..19689
/note="assembly_name:Contig34"
19790..31278
/note="assembly_name:Contig35"
31379..51289
/note="assembly_name:Contig36"
51390..81880
/note="assembly_name:Contig37"
81981..124749
/note="assembly_name:Contig38"
124850..164718
/note="assembly_name:Contig39"
164819..211003
/note="assembly_name:Contig40"

misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature

ORIGIN
Query Match 37.6%; Score 460.4; DB 2; Length 211003;
Best Local Similarity 99.8%; Pred. No. 1.7e-126;
Matches 461; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```



```

QY 763 GACATCTTTACAGGATGATAGCTCAAAAGTTGAGACACATTTGTTAGCACAAACCTGG 822
D 175737 GACATCTTTACAGGATGATAGCTCAAAAGTTGAGACACATTTGTTAGCACAAACCTGG 175678
QY 823 CAGAAAAGAGAAACAGAGCTTCCCTTCAAACTGTTTCCCTGCTTACCATGTTCTACACATC 882
D 175677 CAGAAAAGAGAAACAGAGCTTCCCTTCAAACTGTTTCCCTGCTTACCATGTTCTACACATC 175618
QY 883 AAGTCATGGGGTAATCTCCAAAGTCTTACTTCAAGTCTTACTTCAAGTCTTCCGCAAGACCAATTCGAATGG 942
D 175617 AAGTCATGGGGTAATCTCCAAAGTCTTACTTCAAGTCTTACTTCAAGTCTTCCGCAAGACCAATTCGAATGG 175558
QY 943 TGTGTTTCCATAAAGGCTCCGCAATCTCGTGCAGCTGATTCGACCTTAATTCGAAGC 1002
D 175557 TGTGTTTCCATAAAGGCTCCGCAATCTCGTGCAGCTGATTCGACCTTAATTCGAAGC 175498
QY 1003 CTACACCAAGCCCTTAAGAGTGGAGGATTCATCTGTACAAAGATCACTACATTTACCAG 1062
D 175497 CTACACCAAGCCCTTAAGAGTGGAGGATTCATCTGTACAAAGATCACTACATTTACCAG 175438
QY 1063 GCATTTCAATAATATATCTCCGTTATGGTTCGTGAAGTAACTCGTGAAAGGCCACA 1122
D 175437 GCATTTCAATAATATATCTCCGTTATGGTTCGTGAAGTAACTCGTGAAAGGCCACA 175378
QY 1123 CCTCTGCTCTGAAACACCTGGCAGTGGACATCTCGCTGATCTGTTCTCCATAAT 1182
D 175377 CCTCTGCTCTGAAACACCTGGCAGTGGACATCTCGCTGATCTGTTCTCCATAAT 175318
QY 1183 ATCAAGGCTTCTGAGTGAGCACACAGTGGCTGCTCAATAAAG 1224
D 175317 TTCAGGCTTCTGAGTGAGCACACAGTGGCTGCTCAATAAAG 175276

RESULT 14
AC114618 255925 bp DNA linear HTG 21-OCT-2002
LOCUS Mus musculus clone RP24-86C15, WORKING DRAFT SEQUENCE, 4 ordered
DEFINITION Mus musculus clone RP24-86C15, WORKING DRAFT SEQUENCE, 4 ordered
pieces.
AC114618
VERSION AC114618.4 GI:24182204
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS FULLTOP.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 255925)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
JOURNAL Unpublished
TITLE Mus musculus, clone RP24-86C15
REFERENCE 2 (bases 1 to 255925)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Collangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Jones,C.,
Lander,E., Horton,L., Hume,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczyk,J., Levine,R., Major,J., Marquis,N., Matthews,C.,
MacLean,C., Macdonald,P., Maylor,J., Naylor,J., Nguyen,C., Nicol,R.,
McCarthy,M., McEwan,P., McKernan,K., Melidrim,J., Meneus,L.,
Mihova,T., Miengwa,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
TITLE

```

```

JOURNAL Submitted (10-MAR-2002) Whitehead Institute/MIT Center for Genome
REFERENCE Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS 3 (bases 1 to 255925)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., S.,
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagos,B., Horton,L., Hume,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J.,
Matthews,C., McCarthy,M., Melidrim,J., Meneus,L., Mihova,T.,
Miengwa,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (21-OCT-2002) Whitehead Institute/MIT Center for Genome
REFERENCE Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Oct 21, 2002 this sequence version replaced gi:21431369.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L24322
Center clone name: 86_C15
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 254718 bases at least Q40
Consensus quality: 25230 bases at least Q30
Consensus quality: 25385 bases at least Q20
Insert size: 215000; agarose-fp
Insert size: 255625; sum-of-contigs
Quality coverage: 10.3 in Q20 bases; agarose-fp
Quality coverage: 8.6 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 32897: contig of 32897 bp in length
* 32898 32997: gap of 100 bp
* 32998 76125: contig of 43128 bp in length
* 76126 76225: gap of 100 bp
* 76226 132416: contig of 56191 bp in length
* 132417 132516: gap of 100 bp
* 132517 255925: contig of 123409 bp in length.
* Location/Qualifiers
* 1..255925
* /organism="Mus musculus"
* /mol_type="genomic DNA"
* /db_xref="taxon:10090"
* /clone="RP24-86C15"
* /clone_lib="RPCI-24 Male Mouse BAC"
* 1..32897
* /note="assembly_fragment"
* misc_feature

```



Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Geuregeorgis, E., Geer, K., Hamilton, C., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hughes, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levay, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhewa, L., Loulsegad, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangun, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mathew, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, I., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokelimeh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Popovic, D., Primus, E., Pu, L., L., Plopper, F., Poindexter, A., Popovic, D., Sitter, C.D., Smajs, D., Puzao, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steidle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, I., Walker, B., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willison, R., Wlecyuk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, X., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission  
Unpublished  
2 (bases 1 to 231698)  
2 (bases 1 to 231698)  
Worley, K.C.  
Direct Submission  
Submitted (13-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 231698)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Nov 19, 2002 this sequence version replaced gi:23672467.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GVOQ  
Center clone name: CH230-42M22  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 218837 bases at least 040

Consensus quality: 220614 bases at least Q30  
 Consensus quality: 222122 bases at least Q20  
 Estimated insert size: 227015; sum-of-contigs estimation  
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 2 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 230572: contig of 230572 bp in length  
 \* 230573 230672: gap of unknown length  
 \* 230673 231698: contig of 1026 bp in length.

FEATURES  
 source  
 1. 231698  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-42M22"  
 complement (228283..228997)  
 /note="clone boundary  
 clone\_end:Sp6  
 site:  
 end\_sequence:BH261023"

ORIGIN  
 Query Match 25.2%; Score 308.2; DB 2; Length 231698;  
 Best Local Similarity 80.7%; Pred. No. 8.8e-81;  
 Matches 372; Conservative 0; Mismatches 88; Indels 1; Gaps 1;  
 QY 763 GACATCTTTACAGATGATAGCTCAAAAGTTGAAGACACATTTGTTAGCACAAACCTGG 822  
 Db 136791 GACATCTTTGAGCTGATAGCTCAAAAGCTGAAGTGCATTTCTGTGTAATCCTGG 136850  
 QY 823 CAGAAAAAGAACAGAGCTTCCTTCAAACTGTTCCCTCCCTTACCATGCTACACATC 882  
 Db 136851 CAGCGAAAGAACACAGAGCTTCCTTCAAACTGTTCCCTCCCTTACCATGCTACACATC 136910  
 QY 883 AAGTCCATTTGGGTAACTTCCAGCTTACTTCACTTCAGTCTCGCAAGACCATTCCAAATGG 942  
 Db 136911 AAGGCCATTTGGGGACCTCTCGAGTCCGACTTCCCTTCTCATCAGCACCATTCCAAATGG 136970  
 QY 943 TGTGTTTCCATAAAGGGCTCCGCAATCGCTGGACCTCATTTGGAGACCTAAATCGAAGC 1002  
 Db 136971 TGTGTTTCCATAAAGGACTCTCAGGCTCGCTGGACCTCATTTGGAGACCTAAATCGGAGC 137030  
 QY 1003 CTACACCAAGCTTAAAGGTGGAGGATCTATCTGTACAAGATCACTACATTTACCAG 1062  
 Db 137031 CCACATCAAGCTTTGAGAAGTGGAGGATTTATCTGTCGAAGATCGGTACATTTACCAG 137090  
 QY 1063 GCATTTCAATAATATATCTCGTTATCGGTTCTGTAAGTAACTCGGTGAAGGCCACA 1122  
 Db 137091 TCATTTGATAGTTAGTTTCCATTTATGCTTCTTAACTAGCTTGGGGAAGTCCCA 137150  
 QY 1123 CCTCTGCTCTTGAACACACTGGCACTGGACATCTCGCTTTGGATCTGTTCTCCATAAT 1182  
 Db 137151 CCCACTGCCCTTAAAGACACTGGCTCTGGACATCTTGCCTTGGATCTATCTCCATATG 137210  
 QY 1183 ATCAAGGCTTCTGATGAGACCAACGTAGCGTCCCAATAAA 1223  
 Db 137211 TTAAGAGTTTCTGAGTG-TCCTAACATCTCATCCAATAAA 137250

Search completed: October 14, 2004, 03:42:37  
 Job time : 4732.53 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2004, 00:29:11 ; Search time 48.2954 Seconds  
(without alignments)  
1156.899 Million cell updates/sec

Title: US-10-790-589-2  
Perfect score: 1927  
Sequence: 1 MTAKPLRTVLSLLFFALSGV.....KNHIYQAFHKLRYGPK 354

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 838593 seqs, 157832909 residues

Total number of hits satisfying chosen parameters: 838593

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents AA New:  
1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US05\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1927	100.0	354	US-10-790-589-2	Sequence 2, Appl
2	1293.5	67.1	357	US-10-790-589-4	Sequence 4, Appl
3	1293.5	67.1	361	US-10-170-205E-23316	Sequence 23316, A
4	803	31.3	378	US-10-940-774-8375	Sequence 8375, Ap
5	596.5	31.0	360	PCT-US04-02974-42	Sequence 42, Appl
6	596.5	31.0	360	PCT-US04-12706-30	Sequence 30, Appl
7	596.5	31.0	360	PCT-US04-12706-45	Sequence 45, Appl
8	596.5	31.0	360	US-10-770-668-42	Sequence 42, Appl
9	596.5	31.0	360	US-10-830-828-30	Sequence 30, Appl
10	596.5	31.0	360	US-10-830-828-45	Sequence 45, Appl
11	596.5	31.0	360	US-10-830-828-45	Sequence 45, Appl
12	559	29.0	361	US-10-170-205E-36038	Sequence 36038, A
13	559	29.0	361	US-60-565-632-8103	Sequence 8103, Ap
14	535	27.8	375	PCT-US04-12706-61	Sequence 61, Appl
15	535	27.8	375	US-10-830-828-61	Sequence 61, Appl
16	462.5	24.0	373	PCT-US04-12706-14	Sequence 14, Appl
17	462.5	24.0	373	PCT-US04-12706-38	Sequence 38, Appl
18	462.5	24.0	373	US-10-830-828-14	Sequence 14, Appl
19	462.5	24.0	373	US-10-830-828-38	Sequence 38, Appl
20	462.5	24.0	378	PCT-US04-12706-16	Sequence 16, Appl
21	462.5	24.0	378	US-10-830-828-16	Sequence 16, Appl
22	101.5	5.3	475	PCT-US04-24424-1936	Sequence 1936, Ap
23	101.5	5.3	475	US-10-793-373-26	Sequence 26, Appl
24	101.5	5.3	475	US-10-821-801-499	Sequence 499, App
25	101.5	5.3	475	US-10-913-553-142	Sequence 142, App
26	101.5	5.3	475	US-10-940-774-6067	Sequence 6067, Ap

27	94	4.9	2165	6	US-10-789-400-13	Sequence 13, Appl
28	94	4.9	2165	6	US-10-876-113-29	Sequence 29, Appl
29	94	4.9	2165	6	US-10-876-113-30	Sequence 30, Appl
30	90	4.7	1445	7	US-60-584-405-181	Sequence 181, App
31	89	4.6	1451	6	US-10-170-205E-28165	Sequence 28165, A
32	88.5	4.6	2771	6	US-10-844-295-70	Sequence 70, Appl
33	88	4.6	447	7	US-60-592-978-12635	Sequence 12635, A
34	87.5	4.5	471	7	US-60-592-978-21555	Sequence 21555, A
35	87.5	4.5	1278	1	PCT-US04-07626-51	Sequence 51, Appl
36	87.5	4.5	1279	6	US-10-170-205E-28049	Sequence 28049, A
37	87	4.5	473	7	US-60-592-978-19070	Sequence 19070, A
38	87	4.5	625	6	US-10-917-503-11266	Sequence 11266, A
39	87	4.5	836	6	US-10-425-115-359074	Sequence 359074, A
40	87	4.5	964	6	US-10-917-503-12355	Sequence 12355, A
41	87	4.5	1045	6	US-10-917-503-11431	Sequence 11431, A
42	87	4.5	1094	6	US-10-821-234-1097	Sequence 1097, Ap
43	87	4.5	1798	6	US-10-170-205E-36203	Sequence 36203, A
44	87	4.5	2613	6	US-10-455-772-530	Sequence 530, App
45	87	4.5	2628	6	US-10-455-772-502	Sequence 502, App

## ALIGNMENTS

RESULT 1  
US-10-790-589-2  
; Sequence 2, Application US/10790589  
; GENERAL INFORMATION:  
; APPLICANT: Eastman, Alan R.  
; APPLICANT: Trustees of Dartmouth College  
; TITLE OF INVENTION: DEOXYRIBONUCLEASE II BETA PROTEINS AND CDNAS  
; FILE REFERENCE: DC-0154  
; CURRENT APPLICATION NUMBER: US/10790,589  
; CURRENT FILING DATE: 2004-03-01  
; PRIOR APPLICATION NUMBER: 09/541,840  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: 09/574,942  
; PRIOR FILING DATE: 2000-05-19  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 354  
; TYPE: PRT  
; ORGANISM: Mus sp.  
US-10-790-589-2

	Query Match	100.0%	Score 1927;	DB 6;	Length 354;
	Best Local Similarity	100.0%;	Pred. No. 1.3e-177;		
	Matches 354;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MTAKPLRTVLSLLFFALSGVLTGTPRISCRNEYGEADVNFIFYKLPRKTSKASEAGLQYL	60		
Db	1	MTAKPLRTVLSLLFFALSGVLTGTPRISCRNEYGEADVNFIFYKLPRKTSKASEAGLQYL	60		
QY	61	YLDSTFOTWNSKSLYLINSTRSALGRTLOHLYDTHNSTNDTAYLIYNDGPGSVNYSRQYG	120		
Db	61	YLDSTFOTWNSKSLYLINSTRSALGRTLOHLYDTHNSTNDTAYLIYNDGPGSVNYSRQYG	120		
QY	121	HAKGLLVNRTQGFLLIHSVPKFPVHGVEYPTSGRRYGTGTCITFGYSQFEEIDFOLL	180		
Db	121	HAKGLLVNRTQGFLLIHSVPKFPVHGVEYPTSGRRYGTGTCITFGYSQFEEIDFOLL	180		
QY	181	VLQPNYSCFIPSTTHFWKLIYMPRCANSSSLIKI PVRYLAELHSAQGLNFVHFAKSSFT	240		
Db	181	VLQPNYSCFIPSTTHFWKLIYMPRCANSSSLIKI PVRYLAELHSAQGLNFVHFAKSSFT	240		
QY	241	DDIFGTWIAQKLKTHLLAQTWOKKQELPSNCSLPVHVNIKSIQVTSKYSFSSRODHSK	300		
Db	241	DDIFGTWIAQKLKTHLLAQTWOKKQELPSNCSLPVHVNIKSIQVTSKYSFSSRODHSK	300		
QY	301	WCVSTKSGANRWTCIGDLNRSIHOALRGGGFTCTKNHIYQAFHKLRYGPK	354		

```

Db      301 WCVSIKGSANRWTCIGDLNRSLSHQALRGGGFICTKNHYIYQAFHKLRLRYGFK 354

RESULT 2
US-10-790-589-4
; Sequence 4, Application US/10790589
; GENERAL INFORMATION:
; APPLICANT: Eastman, Alan R.
; APPLICANT: Krieser, Ronald J.
; TITLE OF INVENTION: Deoxyribonuclease II beta proteins and cDNAs
; FILE REFERENCE: DC-0154
; CURRENT APPLICATION NUMBER: US/10790,589
; CURRENT FILING DATE: 2004-03-01
; PRIOR APPLICATION NUMBER: 09/541,840
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 09/574,942
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-790-589-4

Query Match      67.1%; Score 1293.5; DB 6; Length 357;
Best Local Similarity 66.4%; Pred. No. 2.1e-116;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

QY      1 MTKAPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASERAGLYL 60
Db      1 MTKAPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASERAGLYL 60
QY      61 YLDSTRTQWNKSLYLINSTSALGRTHLQHYDTHNS--TNDTAYLIYNDGVPKPNYSRQY 119
Db      61 YLDSTRTQWNKSLYLINSTSALGRTHLQHYDTHNS--TNDTAYLIYNDGVPKPNYSRQY 119
QY      120 GHAKGLLVNWRTOGFWLIHSVPKPPV--HGVEYPTSGRRYGQGTGICITFGYSOFSEIDF 177
Db      120 GHAKGLLVNWRTOGFWLIHSVPKPPV--HGVEYPTSGRRYGQGTGICITFGYSOFSEIDF 177
QY      121 GHTKGLLVNWRVQGFVLIHSIPQPPPIPEEGYDYPPTGRRNGQSGICITPKYNGYEALDS 180
Db      121 GHTKGLLVNWRVQGFVLIHSIPQPPPIPEEGYDYPPTGRRNGQSGICITPKYNGYEALDS 180
QY      178 QLLVLPNIYSCFIPSTFHWKLIYMPKMCANSSSLKIPVRYLAELHSAQGLNFVHFAKSS 237
Db      178 QLLVLPNIYSCFIPSTFHWKLIYMPKMCANSSSLKIPVRYLAELHSAQGLNFVHFAKSS 237
QY      181 QLLVCNPNVYSCSIPATFHQELIHPQLCTFASSEIPGRLLTTLQSAQGGKFLHFAKSD 240
Db      181 QLLVCNPNVYSCSIPATFHQELIHPQLCTFASSEIPGRLLTTLQSAQGGKFLHFAKSD 240
QY      238 FYTDDIFTGWIQAOKLTHLLAQTKWKKQELPSNCSLPYHVNKISIGVTSKSYFSSRQD 297
Db      238 FYTDDIFTGWIQAOKLTHLLAQTKWKKQELPSNCSLPYHVNKISIGVTSKSYFSSRQD 297
QY      241 SFLLDDIFAAMWAQRLKTHLLTETWQRKQELPSNCSLPYHVNKIKALKSRHSYFSSYQD 300
Db      241 SFLLDDIFAAMWAQRLKTHLLTETWQRKQELPSNCSLPYHVNKIKALKSRHSYFSSYQD 300
QY      298 HSKWCVSIKGSANRWTCIGDLNRSLSHQALRGGGFICTKNHYIYQAFHKLRLRYGFK 354
Db      298 HSKWCVSIKGSANRWTCIGDLNRSLSHQALRGGGFICTKNHYIYQAFHKLRLRYGFK 354

RESULT 3
US-10-170-205E-23316
; Sequence 23316, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 23316
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-23316

```

```

Query Match      67.1%; Score 1293.5; DB 6; Length 361;
Best Local Similarity 66.4%; Pred. No. 2.1e-116;
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;

QY      1 MTKAPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASERAGLYL 60
Db      1 MTKAPLRTVLSLFFALSGVLGTPETISCRNEYGEAVDWFIFYKLPKRTSKASERAGLYL 60
QY      61 YLDSTRTQWNKSLYLINSTSALGRTHLQHYDTHNS--TNDTAYLIYNDGVPKPNYSRQY 119
Db      61 YLDSTRTQWNKSLYLINSTSALGRTHLQHYDTHNS--TNDTAYLIYNDGVPKPNYSRQY 119
QY      120 GHAKGLLVNWRTOGFWLIHSVPKPPV--HGVEYPTSGRRYGQGTGICITFGYSOFSEIDF 177
Db      120 GHAKGLLVNWRTOGFWLIHSVPKPPV--HGVEYPTSGRRYGQGTGICITFGYSOFSEIDF 177
QY      125 GHTKGLLVNWRVQGFVLIHSIPQPPPIPEEGYDYPPTGRRNGQSGICITPKYNGYEALDS 184
Db      125 GHTKGLLVNWRVQGFVLIHSIPQPPPIPEEGYDYPPTGRRNGQSGICITPKYNGYEALDS 184
QY      178 QLLVLPNIYSCFIPSTFHWKLIYMPKMCANSSSLKIPVRYLAELHSAQGLNFVHFAKSS 237
Db      178 QLLVLPNIYSCFIPSTFHWKLIYMPKMCANSSSLKIPVRYLAELHSAQGLNFVHFAKSS 237
QY      185 QLLVCNPNVYSCSIPATFHQELIHPQLCTFASSEIPGRLLTTLQSAQGGKFLHFAKSD 244
Db      185 QLLVCNPNVYSCSIPATFHQELIHPQLCTFASSEIPGRLLTTLQSAQGGKFLHFAKSD 244
QY      238 FYTDDIFTGWIQAOKLTHLLAQTKWKKQELPSNCSLPYHVNKISIGVTSKSYFSSRQD 297
Db      238 FYTDDIFTGWIQAOKLTHLLAQTKWKKQELPSNCSLPYHVNKISIGVTSKSYFSSRQD 297
QY      245 SFLLDDIFAAMWAQRLKTHLLTETWQRKQELPSNCSLPYHVNKIKALKSRHSYFSSYQD 304
Db      245 SFLLDDIFAAMWAQRLKTHLLTETWQRKQELPSNCSLPYHVNKIKALKSRHSYFSSYQD 304
QY      298 HSKWCVSIKGSANRWTCIGDLNRSLSHQALRGGGFICTKNHYIYQAFHKLRLRYGFK 354
Db      298 HSKWCVSIKGSANRWTCIGDLNRSLSHQALRGGGFICTKNHYIYQAFHKLRLRYGFK 354

RESULT 4
US-10-940-774-8375
; Sequence 8375, Application US/10940774
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/10/940,774
; CURRENT FILING DATE: 2004-09-15
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8375
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Human
US-10-940-774-8375

Query Match      31.3%; Score 603; DB 6; Length 378;
Best Local Similarity 37.6%; Pred. No. 1.1e-49;
Matches 135; Conservative 59; Mismatches 147; Indels 18; Gaps 8;

QY      5 PLRTVLSLFFALSGVLGTP--HISCRNEYGEAVDWFIFYKLPK-RTSKASERAGLYL 61
Db      5 PLRTVLSLFFALSGVLGTP--HISCRNEYGEAVDWFIFYKLPK-RTSKASERAGLYL 61
QY      15 PIAMTPLL---LAALLCVPAAGALTCTGDSGQPVDFVYKLPALRGSGGAQGLQYKY 71
Db      15 PIAMTPLL---LAALLCVPAAGALTCTGDSGQPVDFVYKLPALRGSGGAQGLQYKY 71
QY      62 LDSTRTQWNKSLYLINSTSALGRTHLQHYDTHNSTNDTAYLIYND--GVPKPNYSRQY 119
Db      62 LDSTRTQWNKSLYLINSTSALGRTHLQHYDTHNSTNDTAYLIYND--GVPKPNYSRQY 119
QY      72 LDSSGGWRDGRALINSPGAVGRSLQPLY--RSNTSQAFLLYNDQPPQSKAQDSSMR 129
Db      72 LDSSGGWRDGRALINSPGAVGRSLQPLY--RSNTSQAFLLYNDQPPQSKAQDSSMR 129
QY      120 GHAKGLLVNWRTOGFWLIHSVPKPPV--HGVEYPTSGRRYGQGTGICITFGYSOFSEID 176
Db      120 GHAKGLLVNWRTOGFWLIHSVPKPPV--HGVEYPTSGRRYGQGTGICITFGYSOFSEID 176
QY      130 GHTKGLLVLDHGGFWLVHVPNPPPPASAAAYSWPHSACTYGTLLCVSPFPFAQFKMG 189
Db      130 GHTKGLLVLDHGGFWLVHVPNPPPPASAAAYSWPHSACTYGTLLCVSPFPFAQFKMG 189
QY      177 FOLLVLPNIYSCFIPSTFHWKLIYMPKMCANSSSLKIPVRYLAELHSAQGLNFVHFAKS 236
Db      177 FOLLVLPNIYSCFIPSTFHWKLIYMPKMCANSSSLKIPVRYLAELHSAQGLNFVHFAKS 236
QY      190 KQLTYTPWVINYQLEGIFAQEPFDLENVVKGHVSQEPWNSSITLISQAGVQSPAKF 249
Db      190 KQLTYTPWVINYQLEGIFAQEPFDLENVVKGHVSQEPWNSSITLISQAGVQSPAKF 249

```

QY 237 SFYTDIDFTGIAQKLTLLAQTWQKKQELPNSCLPYHVYVNIKSIGVTSKS--YFSS 294  
 Db 250 SKFGDDLYSGWLAALGTNLQVFWHTVGLPNSCSDIWOVLNVNQLAFPGPAGSFSNS 309  
 QY 295 RQDHSKWCVSIGKANRWTCIGDLNLSHQAALGGGFTCTKNHYIYOAFHKLILRYGFC 353  
 Db 310 TEDHSKWCVSIPKGP---WTCVGMNMQEQRGGGTLCALPALMKAFQPLVKNYQPC 365

RESULT 5

PCT-US04-02974-42  
 ; Sequence 42, Application PC/TUS0402974  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wright, Susan C.  
 ; APPLICANT: Larrick, James W.  
 ; APPLICANT: Nock, Steffen R.  
 ; APPLICANT: Wilson, David S.  
 ; TITLE OF INVENTION: Cell-Killing Molecules and Methods of Use Thereof  
 ; FILE REFERENCE: ABSALUS-08804  
 ; CURRENT APPLICATION NUMBER: PCT/US04/02974  
 ; CURRENT FILING DATE: 2004-02-02  
 ; NUMBER OF SEQ ID NOS: 81  
 ; SOFTWARE: Patent in version 3.2  
 ; SEQ ID NO 42  
 ; LENGTH: 360  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; PCT-US04-02974-42

Query Match 31.0%; Score 596.5; DB 1; Length 360;  
 Best Local Similarity 38.0%; Pred. No. 4.5e-49;  
 Matches 132; Conservative 56; Mismatches 144; Indels 15; Gaps 7;

QY 17 LSGVLGTP--EISCRNEYGEAVDMFIFYKLPK-RTSKASEEAGLOLYLDSTROTWNKSL 73  
 Db 6 LAALLCVPAGALTCYGDGSGQPDVFWVYKLPALRGSGEAAQRLQKYLDESSGGWRDGR 65  
 QY 74 YLINSRSLGRTLOHLYDTHNSTNDTAVLIYND--GVPGSVNYSRQYGHAKGLLVNRT 131  
 Db 66 ALINSEGAAGVRSLOPLY--RSNTSOLAFLYNDQPPQSKAQDSSMRGHTKGVLLDHD 123  
 QY 132 QGFWLHSPKPPPP---VHGYEPTSGRRYGTGICITFGYSQFEIDFQLLVLPNIYS 188  
 Db 124 GGFWLHSPVNPFPFPASSAAYSWPHSACTYGTLLCVSPFPAQFSKMGKQLTYTPWVYN 183  
 QY 189 CFIPSTFHWKLIYMPRMCANSSSLKIPVRYLAELHSAQGLNFVHFAKSSFYTDDIFTGMI 248  
 Db 184 YLEGIFAQEFDPDENNVVGHVSPFPWNSITLTSQAGAVFQSFQSKFGDGLYSGWL 243  
 QY 249 AQKLKTHLLAQTWQKKQELPNSCLPYHVYVNIKSIGVTSKS--YFSSRQDHSKWCVSIK 306  
 Db 244 AAALGTNLQVFWHTVGLPNSCSDIWOVLNVNQLAFPGPAGSFSNSTEDHSKWCVSIPK 303  
 QY 307 GSNRWTCIGDLNLSHQAALGGGFTCTKNHYIYOAFHKLILRYGFC 353  
 Db 304 GP---WTCVGMNMQEQRGGGTLCALPALMKAFQPLVKNYQPC 347

RESULT 6

PCT-US04-12706-30  
 ; Sequence 30, Application PC/TUS0412706  
 ; GENERAL INFORMATION:  
 ; APPLICANT: The Regents of the University of Colorado  
 ; APPLICANT: Parrish, Jay  
 ; APPLICANT: Xue, Ding  
 ; TITLE OF INVENTION: NUCLEASE GENES AND THEIR USES  
 ; FILE REFERENCE: 420295  
 ; CURRENT APPLICATION NUMBER: PCT/US04/12706  
 ; CURRENT FILING DATE: 2004-04-29  
 ; PRIOR APPLICATION NUMBER: US 60/498,065  
 ; PRIOR FILING DATE: 2003-08-26  
 ; PRIOR APPLICATION NUMBER: US 60/465,086  
 ; PRIOR FILING DATE: 2003-04-23

; NUMBER OF SEQ ID NOS: 62  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 30  
 ; LENGTH: 360  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; PCT-US04-12706-30

Query Match 31.0%; Score 596.5; DB 1; Length 360;  
 Best Local Similarity 38.0%; Pred. No. 4.5e-49;  
 Matches 132; Conservative 56; Mismatches 144; Indels 15; Gaps 7;

QY 17 LSGVLGTP--EISCRNEYGEAVDMFIFYKLPK-RTSKASEEAGLOLYLDSTROTWNKSL 73  
 Db 6 LAALLCVPAGALTCYGDGSGQPDVFWVYKLPALRGSGEAAQRLQKYLDESSGGWRDGR 65  
 QY 74 YLINSRSLGRTLOHLYDTHNSTNDTAVLIYND--GVPGSVNYSRQYGHAKGLLVNRT 131  
 Db 66 ALINSEGAAGVRSLOPLY--RSNTSOLAFLYNDQPPQSKAQDSSMRGHTKGVLLDHD 123  
 QY 132 QGFWLHSPKPPPP---VHGYEPTSGRRYGTGICITFGYSQFEIDFQLLVLPNIYS 188  
 Db 124 GGFWLHSPVNPFPFPASSAAYSWPHSACTYGTLLCVSPFPAQFSKMGKQLTYTPWVYN 183  
 QY 189 CFIPSTFHWKLIYMPRMCANSSSLKIPVRYLAELHSAQGLNFVHFAKSSFYTDDIFTGMI 248  
 Db 184 YLEGIFAQEFDPDENNVVGHVSPFPWNSITLTSQAGAVFQSFQSKFGDGLYSGWL 243  
 QY 249 AQKLKTHLLAQTWQKKQELPNSCLPYHVYVNIKSIGVTSKS--YFSSRQDHSKWCVSIK 306  
 Db 244 AAALGTNLQVFWHTVGLPNSCSDIWOVLNVNQLAFPGPAGSFSNSTEDHSKWCVSIPK 303  
 QY 307 GSNRWTCIGDLNLSHQAALGGGFTCTKNHYIYOAFHKLILRYGFC 353  
 Db 304 GP---WTCVGMNMQEQRGGGTLCALPALMKAFQPLVKNYQPC 347

RESULT 7

PCT-US04-12706-45  
 ; Sequence 45, Application PC/TUS0412706  
 ; GENERAL INFORMATION:  
 ; APPLICANT: The Regents of the University of Colorado  
 ; APPLICANT: Parrish, Jay  
 ; APPLICANT: Xue, Ding  
 ; TITLE OF INVENTION: NUCLEASE GENES AND THEIR USES  
 ; FILE REFERENCE: 420295  
 ; CURRENT APPLICATION NUMBER: PCT/US04/12706  
 ; CURRENT FILING DATE: 2004-04-29  
 ; PRIOR APPLICATION NUMBER: US 60/498,065  
 ; PRIOR FILING DATE: 2003-08-26  
 ; PRIOR APPLICATION NUMBER: US 60/465,086  
 ; PRIOR FILING DATE: 2003-04-23  
 ; NUMBER OF SEQ ID NOS: 62  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 45  
 ; LENGTH: 360  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: MISC FEATURE  
 ; LOCATION: (1)-(360)  
 ; OTHER INFORMATION: DNaseII (Human CRN-6 homolog)  
 ; PCT-US04-12706-45

Query Match 31.0%; Score 596.5; DB 1; Length 360;  
 Best Local Similarity 38.0%; Pred. No. 4.5e-49;  
 Matches 132; Conservative 56; Mismatches 144; Indels 15; Gaps 7;

QY 17 LSGVLGTP--EISCRNEYGEAVDMFIFYKLPK-RTSKASEEAGLOLYLDSTROTWNKSL 73  
 Db 6 LAALLCVPAGALTCYGDGSGQPDVFWVYKLPALRGSGEAAQRLQKYLDESSGGWRDGR 65  
 QY 74 YLINSRSLGRTLOHLYDTHNSTNDTAVLIYND--GVPGSVNYSRQYGHAKGLLVNRT 131

```
Db 66 ALINSPGAVGRSLQPLY--RSNTSQLAFLLYNDQPPQSKAQDSSMRGHTKGVLILLDHD 123
Qy 132 QGFWLIHSPKFP--VHGVEYPTSGRRYGTGICITFGYSQPEEDFOLLVLQPNYIS 188
Db 124 GGFVLHVSVPNFP--PASSAAYSWPHSACTYGTQLLCVSPFAQSKMGKQLTYTPWVYN 183
Qy 189 CFIPSTHFKLIYPMRCANSSSLKIPVRYLAELHSAQGLNFVHFAKSSFTDDIFTGWI 248
Db 184 YQLEGIFAQEPDLENVVKGHVVSQEPWNSITLTSQAGAVFQSFAPKSKFGDDLYSGWL 243
Qy 249 AQKLKTHLLAQTWQKKQELPNSCLSPHYVNIKSIQVTSKS--YFSSRQDHSKWCVSIK 306
Db 244 AALGTNLQVFWHTKVTGILPNSCSDIWQVLNVNQIAFPAGPAGPSFNSTEDHSKWCVSIPK 303
Qy 307 GSANRWTCIGDLNLSHQAALRGGGFICTKNHYIYQAFHKLRYGFC 353
Db 304 GP---WTCVGMNRRNQEGEORGGGTLCALPALWKAQFPLVKNYQPC 347
```

## RESULT 8

US-10-770-668-42

```
; Sequence 42, Application US/10770668
; GENERAL INFORMATION:
; APPLICANT: Wright, Susan C.
; APPLICANT: Larick, James W.
; APPLICANT: Nock, Steffen R.
; APPLICANT: Wilson, David S.
; TITLE OF INVENTION: Cell-killing Molecules and Methods of Use Thereof
; FILE REFERENCE: ABSALUS-08602
; CURRENT APPLICATION NUMBER: US/10/770,668
; CURRENT FILING DATE: 2004-02-02
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 42
; LENGTH: 360
; TYPE: PRP
; ORGANISM: Homo sapiens
US-10-770-668-42
```

```
Query Match 31.0%; Score 596.5; DB 6; Length 360;
Best Local Similarity 38.0%; Pred. No. 4.5e-49;
Matches 132; Conservative 56; Mismatches 144; Indels 15; Gaps 7;

Qy 17 LSGVLGTP--EISCRNEYGEAVDMFIFKLPK-RTSKASEAGLOYLVDSTROTWNKSL 73
Db 6 LALLCVPAGALTTCYGDGQPDVWFVVKLPALRGSGEAAQRLQYKYLDESSGGWRDGR 65
```

```
Qy 74 YLINSTRSALGRITLQHLTDHSTNDTAYLIYND--GVPGSVNYSRQYGHAKGLLVNRT 131
Db 66 ALINSPGAVGRSLQPLY--RSNTSQLAFLLYNDQPPQSKAQDSSMRGHTKGVLILLDHD 123
Qy 132 QGFWLIHSPKFP--VHGVEYPTSGRRYGTGICITFGYSQPEEDFOLLVLQPNYIS 188
Db 124 GGFVLHVSVPNFP--PASSAAYSWPHSACTYGTQLLCVSPFAQSKMGKQLTYTPWVYN 183
Qy 189 CFIPSTHFKLIYPMRCANSSSLKIPVRYLAELHSAQGLNFVHFAKSSFTDDIFTGWI 248
Db 184 YQLEGIFAQEPDLENVVKGHVVSQEPWNSITLTSQAGAVFQSFAPKSKFGDDLYSGWL 243
Qy 249 AQKLKTHLLAQTWQKKQELPNSCLSPHYVNIKSIQVTSKS--YFSSRQDHSKWCVSIK 306
Db 244 AALGTNLQVFWHTKVTGILPNSCSDIWQVLNVNQIAFPAGPAGPSFNSTEDHSKWCVSIPK 303
Qy 307 GSANRWTCIGDLNLSHQAALRGGGFICTKNHYIYQAFHKLRYGFC 353
Db 304 GP---WTCVGMNRRNQEGEORGGGTLCALPALWKAQFPLVKNYQPC 347
```

## RESULT 9

US-10-830-828-30

```
; Sequence 30, Application US/10830828
; GENERAL INFORMATION:
```

```
; APPLICANT: The Regents of the University of Colorado
; APPLICANT: Parrish, Jay
; APPLICANT: Xue, Ding
; TITLE OF INVENTION: NUCLEASE GENES AND THEIR USES
; FILE REFERENCE: 420295
; CURRENT APPLICATION NUMBER: US/10/830,828
; CURRENT FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/498,065
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: US 60/465,086
; PRIOR FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 30
; LENGTH: 360
; TYPE: PRP
; ORGANISM: Homo sapiens
US-10-830-828-30
```

```
Query Match 31.0%; Score 596.5; DB 6; Length 360;
Best Local Similarity 38.0%; Pred. No. 4.5e-49;
Matches 132; Conservative 56; Mismatches 144; Indels 15; Gaps 7;
```

```
Qy 17 LSGVLGTP--EISCRNEYGEAVDMFIFKLPK-RTSKASEAGLOYLVDSTROTWNKSL 73
Db 6 LALLCVPAGALTTCYGDGQPDVWFVVKLPALRGSGEAAQRLQYKYLDESSGGWRDGR 65
Qy 74 YLINSTRSALGRITLQHLTDHSTNDTAYLIYND--GVPGSVNYSRQYGHAKGLLVNRT 131
Db 66 ALINSPGAVGRSLQPLY--RSNTSQLAFLLYNDQPPQSKAQDSSMRGHTKGVLILLDHD 123
Qy 132 QGFWLIHSPKFP--VHGVEYPTSGRRYGTGICITFGYSQPEEDFOLLVLQPNYIS 188
Db 124 GGFVLHVSVPNFP--PASSAAYSWPHSACTYGTQLLCVSPFAQSKMGKQLTYTPWVYN 183
Qy 189 CFIPSTHFKLIYPMRCANSSSLKIPVRYLAELHSAQGLNFVHFAKSSFTDDIFTGWI 248
Db 184 YQLEGIFAQEPDLENVVKGHVVSQEPWNSITLTSQAGAVFQSFAPKSKFGDDLYSGWL 243
Qy 249 AQKLKTHLLAQTWQKKQELPNSCLSPHYVNIKSIQVTSKS--YFSSRQDHSKWCVSIK 306
Db 244 AALGTNLQVFWHTKVTGILPNSCSDIWQVLNVNQIAFPAGPAGPSFNSTEDHSKWCVSIPK 303
Qy 307 GSANRWTCIGDLNLSHQAALRGGGFICTKNHYIYQAFHKLRYGFC 353
Db 304 GP---WTCVGMNRRNQEGEORGGGTLCALPALWKAQFPLVKNYQPC 347
```

## RESULT 10

US-10-830-828-45

```
; Sequence 45, Application US/10830828
; GENERAL INFORMATION:
; APPLICANT: The Regents of the University of Colorado
; APPLICANT: Parrish, Jay
; APPLICANT: Xue, Ding
; TITLE OF INVENTION: NUCLEASE GENES AND THEIR USES
; FILE REFERENCE: 420295
; CURRENT APPLICATION NUMBER: US/10/830,828
; CURRENT FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/498,065
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: US 60/465,086
; PRIOR FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 45
; LENGTH: 360
; TYPE: PRP
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(360)
; OTHER INFORMATION: DnaaseII (Human CRN-6 homolog)
```



```
; TITLE OF INVENTION: Compositions thereof
; FILE REFERENCE: 38-21 (53403) C
; CURRENT APPLICATION NUMBER: US/60/579,062
; CURRENT FILING DATE: 2004-06-11
; NUMBER OF SEQ ID NOS: 41445
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8103
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Diabrotica virgifera
US-60-579-062-8103

Query Match 29.0%; Score 559; DB 7; Length 361;
Best Local Similarity 35.9%; Pred. No. 1.9e-45;
Matches 122; Conservative 65; Mismatches 143; Indels 10; Gaps 9;

QY 12 LFFALSGVLG-TPEISCRNEYGEAVDFWFIYKLPKRTSKAS-BEAGLQYLYLDSTR-OT 68
DB 6 LFFCLYFIAGRTNALSCDENRVDWFAVKIPIQSGRGDIKAGLGHVLTSTNSST 65

QY 69 WNKSLYLINSTRSALGRTLOHLYDTHNSTNDTAYLIYNDGVPGSVNYSRQYGHAKGLLYW 128
DB 66 WTFKRSIGDSSSLIGNTLHDFY--YNRTK-LSYILYNDEPNHPK-SAAKGHTKGVVIS 121

QY 129 NRTQGFWLHVSVPKPPV-HGVEYPTSGRRYGTGICITFGYSQPEEIDFOLLVLPNIY 187
DB 122 NAAGGYWLHVSVPKPPVPEPVGNSYSPFTGKVGQTLFVTMLNINAVGLQYQNPPIY 181

QY 188 SCFIPSTPHWKLIVPRMCANSSSLKI PVRYLAELHSAQGLNFVHFASKSFYTDIDFTGW 247
DB 182 LQOVLSDIKTRVDPDLASAAQNVVKSAPFYVAQIYSKKVAFTSFAKHKFNQELYEDL 241

QY 248 IAOKLKTHLLAQTWOKKQELPSNCSLPYHVNIKSIGVTSKSY-FSSRDHSHKWCVSIX 306
DB 242 VAPYKQDLVVEVETWPNAGRLHNSGRVYKVNIOKINMNVVNSFNTTSDHSHKAVT-S 300

QY 307 GSANFTWTCIGDLNRSIHQALRGSGFICTKNHYIYQAFHKL 346
DB 301 NSSSWSCVGDINRAQOHLERGCGTVCLQNKILSTNFHKL 340

RESULT 14
PCT-US04-12706-61
; Sequence 61, Application PC/TUS0412706
; GENERAL INFORMATION:
; APPLICANT: The Regents of the University of Colorado
; APPLICANT: Parrish, Jay
; APPLICANT: Xue, Ding
; TITLE OF INVENTION: NUCLEASE GENES AND THEIR USES
; FILE REFERENCE: 420295
; CURRENT APPLICATION NUMBER: PCT/US04/12706
; CURRENT FILING DATE: 2004-04-29
; PRIOR APPLICATION NUMBER: US 60/498,065
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: US 60/465,086
; PRIOR FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)-(375)
; OTHER INFORMATION: NUC-1
PCT-US04-12706-61

Query Match 27.8%; Score 535; DB 1; Length 375;
Best Local Similarity 34.0%; Pred. No. 4.1e-43;
Matches 123; Conservative 59; Mismatches 148; Indels 32; Gaps 8;

QY 11 SLLFFALSGVLGT-PEISCRNEYGEAVDFWFIYKLPKRTSKASE---EAGLQYLYLDSTR 66
DB 7 AVLIFFLLGVSTQYAAFSCKDQSGNDVDFWFAVYKMPTEKDDGSGVTGLAGGVAVIYVDVVK 66

QY 67 Q-TWNSKLYLINSTRSALGRTLOHLYDTHNSTNDTAYLIYNDGVPGS-----112
DB 67 KGTLTPSAKTLDNDQAIAYTLQOYDKQND-KTIFHWYNDPWPWGSKSTSGIKLEELIS 125

QY 113 -----VNYSRQ-----YGHAKGLLVNRTQGFWLHVSVPKPPVHGVEYPTSGRRYGT 161
DB 126 NRVSNTYTHEDDSTSTAFGHTKGTIFPDGTSGLVWLVHVSPLFPNFTKVEYFVSGHDYGT 185

QY 162 GICITFGYSQPEEIDFOLLVLPNIYSCFIPSTPHWKLIVNPRMCANSSSLKIPVRYLA 221
DB 162 GICITFGYSQPEEIDFOLLVLPNIYSCFIPSTPHWKLIVNPRMCANSSSLKIPVRYLA 221

Query Match 27.8%; Score 535; DB 6; Length 375;
Best Local Similarity 34.0%; Pred. No. 4.1e-43;
Matches 123; Conservative 59; Mismatches 148; Indels 32; Gaps 8;

QY 11 SLLFFALSGVLGT-PEISCRNEYGEAVDFWFIYKLPKRTSKASE---EAGLQYLYLDSTR 66
DB 7 AVLIFFLLGVSTQYAAFSCKDQSGNDVDFWFAVYKMPTEKDDGSGVTGLAGGVAVIYVDVVK 66

QY 67 Q-TWNSKLYLINSTRSALGRTLOHLYDTHNSTNDTAYLIYNDGVPGS-----112
DB 67 KGTLTPSAKTLDNDQAIAYTLQOYDKQND-KTIFHWYNDPWPWGSKSTSGIKLEELIS 125

QY 113 -----VNYSRQ-----YGHAKGLLVNRTQGFWLHVSVPKPPVHGVEYPTSGRRYGT 161
DB 126 NRVSNTYTHEDDSTSTAFGHTKGTIFPDGTSGLVWLVHVSPLFPNFTKVEYFVSGHDYGT 185

QY 162 GICITFGYSQPEEIDFOLLVLPNIYSCFIPSTPHWKLIVNPRMCANSSSLKIPVRYLA 221
DB 162 GICITFGYSQPEEIDFOLLVLPNIYSCFIPSTPHWKLIVNPRMCANSSSLKIPVRYLA 221
```



Db	196	MLCMTFKYAQLKSIQTOLFFNRPNIYSSNLPNMAADNADLAKAIAGOYKQGPFSVIE	245
Qy	222	LHSAQGLNFVHFAXSFEYDDIFTGWIAQKLKTHLLAQTWOKKQELPSNCSLPYHVYNI	281
Db	246	LETWAGISFTINFAXSKEFNADLYDTLVAPTLLKTDLVVETW-RRGSEIPLDCKLTYHANDA	304
Qy	282	KSIGVTSKSYFSSRQDHSKWCVSIKGSANRWTCIGDLNRSIHQALRGGGFICTKNHYIQ	341
Db	305	LSIHVGSTTAFSYTKDHSKMAHS-ADWTKPWVCIGDINRWTSQYVRGGGTTCISSSFLWK	363
Qy	342	AP	343
Db	364	AY	365

Search completed: October 14, 2004, 00:43:31  
Job time : 50.2954 secs

***This Page Blank (uspio)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 23:51:40 ; Search time 56.7384 Seconds  
(without alignments)  
1985.252 Million cell updates/sec

Title: US-10-790-589-4  
Perfect score: 1933  
Sequence: 1 MMARLLRTSFALLFLGLFGV.....QNWQIYQAFQGHVLYESCK 357

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_25.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1933	100.0	357	4 Q9NQW3	Q9NQW3 homo sapien
2	1933	100.0	361	4 Q8WZ79	Q8WZ79 homo sapien
3	1297.5	67.1	356	11 Q9QZK9	Q9QZK9 rattus norv
4	1293.5	66.9	354	11 Q9QY48	Q9QY48 mus musculus
5	1290.5	66.8	354	11 Q8C589	Q8C589 mus musculus
6	835	43.2	153	4 Q8WZ80	Q8WZ80 homo sapien
7	762	39.4	375	12 Q90759	Q90759 fowlpox vir
8	560	25.0	350	11 Q9QZK8	Q9QZK8 rattus norv
9	511	26.4	332	12 Q9J5H0	Q9J5H0 fowlpox vir
10	449	23.2	366	5 Q9VED8	Q9VED8 drosophila
11	209	10.8	338	5 Q9BJL8	Q9BJL8 trichinella
12	191.5	9.9	339	5 Q8GQP5	Q8GQP5 trichinella
13	127	6.6	342	5 Q8GQP4	Q8GQP4 trichinella
14	117	6.1	344	5 Q27073	Q27073 trichinella
15	104	5.4	315	5 Q27076	Q27076 trichinella
16	102.5	5.3	698	16 Q823U3	Q823U3 chlamydomphi

17	98.5	5.1	391	16 Q87ZU7	Q87ZU7 pseudomonas
18	97.5	5.0	697	16 Q84355	Q84355 chlamydia t
19	96.5	5.0	704	16 Q9K1Z6	Q9K1Z6 chlamydia p
20	95.5	4.9	308	16 Q9A122	Q9A122 streptococc
21	95.5	4.9	308	16 Q878E6	Q878E6 streptococc
22	95.5	4.9	328	16 Q8K8C3	Q8K8C3 streptococc
23	95.5	4.9	970	5 Q9NEH2	Q9NEH2 leishmania
24	95	4.9	912	16 Q7VGK4	Q7VGK4 helicobacte
25	94.5	4.9	308	16 Q8P240	Q8P240 streptococc
26	94.5	4.9	349	9 Q9G044	Q9G044 bacterioph
27	94.5	4.9	349	16 Q932C3	Q932C3 staphylococ
28	94.5	4.9	2295	5 Q9TV98	Q9TV98 plasmodium
29	93.5	4.8	305	16 Q8A6C7	Q8A6C7 bacteroides
30	93.5	4.8	394	16 Q9PMP4	Q9PMP4 campylobact
31	93.5	4.8	913	13 Q90XB9	Q90XB9 gallus gall
32	93	4.8	466	13 Q90XC4	Q90XC4 poephila gu
33	93	4.8	526	16 Q9PPR1	Q9PPR1 ureaplasma
34	93	4.8	2472	4 Q14214	Q14214 homo sapien
35	92	4.8	641	16 Q8X5E6	Q8X5E6 escherichia
36	92	4.8	641	16 Q8X3B3	Q8X3B3 escherichia
37	92	4.8	1051	4 Q81XU4	Q81XU4 homo sapien
38	91.5	4.7	276	5 Q962V2	Q962V2 trichinella
39	91.5	4.7	930	16 Q7VCX6	Q7VCX6 prochloroco
40	91.5	4.7	984	16 Q8A5H5	Q8A5H5 bacteroides
41	91	4.7	395	7 Q8HWE7	Q8HWE7 mus musculu
42	91	4.7	395	7 Q8HWA5	Q8HWA5 mus musculu
43	91	4.7	395	7 Q8HWA4	Q8HWA4 mus musculu
44	91	4.7	433	6 Q8M1L0	Q8M1L0 oryctolagus
45	91	4.7	433	6 Q8MHY7	Q8MHY7 oryctolagus

ALIGNMENTS

RESULT 1  
Q9NQW3

ID Q9NQW3 PRELIMINARY; PRT; 357 AA.  
AC Q9NQW3;  
DT 01-OCT-2000 (TREMREL. 15, Created)  
DT 01-OCT-2000 (TREMREL. 15, Last sequence update)  
DT 01-JUN-2003 (TREMREL. 24, Last annotation update)  
DE Deoxyribonuclease II beta.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21272509; PubMed=11376952;  
RA Krieger R.J., Maclea K.S., Park J.P., Eastman A.;  
RT "The cloning, genomic structure, localization, and expression of human  
RT deoxyribonuclease II beta."  
RL Gene 269:205-216(2001).  
DR EMBL; AF274571; AAF76893.1; "  
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.  
DR GO; GO:0006259; P:DNA metabolism; IEA.  
DR InterPro; IPR004947; DNaase\_II.  
DR Pfam; PF03265; DNaase II; 1.  
SQ SEQUENCE 357 AA; 41197 MW; 6565D2036BD20955 CRC64;

Query Match 100.0%; Score 1933; DB 4; Length 357;  
Best Local Similarity 100.0%; Pred. No. 1.1e-171;  
Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MMARLLRTSFALLFLGLFGVGLGAATISCRNEEGKAVDMFTFYKLPKRONKESGTEGLYL	60
Db	1	MMARLLRTSFALLFLGLFGVGLGAATISCRNEEGKAVDMFTFYKLPKRONKESGTEGLYL	60
QY	61	YLDSTTSWRKSEQLMNDTKSVLGRITLQQLYEAVASKNNTAYLIYNDGVPKPNYSKY	120
Db	61	YLDSTTSWRKSEQLMNDTKSVLGRITLQQLYEAVASKNNTAYLIYNDGVPKPNYSKY	120
QY	121	CHTKGLLLNWRVQGFLLIHSIPQFPPEEGYDPTGRRNGQSGICITFKYNQVEAIDS	180

```
Db 121 GHTKGLLLNVRVQGFLLIHSIPQPIPEBGIDYPTGRRNGSGGICITFKYQYEAIDS 180
QY 181 QLLVCNPNVYSCSIPATFHQELIHPQLCTRASSEIPGRLLTTLQSAQGGKFLHFAKSD 240
Db 181 QLLVCNPNVYSCSIPATFHQELIHPQLCTRASSEIPGRLLTTLQSAQGGKFLHFAKSD 240
QY 241 SFDDDDFAAWMAQRLKTHLLTETWORKRQBLPNSCLPYHYVNIKAIKLSRHSYFSSYQD 300
Db 241 SFDDDDFAAWMAQRLKTHLLTETWORKRQBLPNSCLPYHYVNIKAIKLSRHSYFSSYQD 300
QY 301 HAKWCISQKGTNRWTCIGDLNRSPhQAfrSGGfICTQNWQIYQAFQGLVLYYESCK 357
Db 301 HAKWCISQKGTNRWTCIGDLNRSPhQAfrSGGfICTQNWQIYQAFQGLVLYYESCK 357

RESULT 2
Q8WZ79
ID Q8WZ79 PRELIMINARY; PRT; 361 AA.
AC Q8WZ79;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Endonuclease DLAD.
GN DLAD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21556924; PubMed=11700027;
RA ShioKawa D., Tanuma S.I.;
RT "Isolation and Characterization of the DLAD/DIAD Genes, Which Lie
RL Biochem. Biophys. Res. Commun. 288:1119-1128(2001).
DR EMBL; AF334460; AAL34449.1; -.
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR004947; DNase_II.
DR Pfam; PF03265; DNase_II; 1.
SQ SEQUENCE 361 AA; 41712 MW; C6FDD3DF58F62CAC0 CRC64;

Query Match 100.0%; Score 1933; DB 4; Length 361;
Best Local Similarity 100.0%; Pred. No. 1.1e-171; Indels 0; Gaps 0;
Matches 357; Conservative 0; Mismatches 0;

QY 1 MMARLLTSPALLFLGLFGVLGAATISCRNEEGKAVDWFYFKLPKQNKESGETGLEYL 60
Db 5 MMARLLTSPALLFLGLFGVLGAATISCRNEEGKAVDWFYFKLPKQNKESGETGLEYL 64
QY 61 YLDSTTRSWRKSEQLMNDTKSVLGRITLQQLYEAVASKNNYALYINDGPKPNYSRKY 120
Db 65 YLDSTTRSWRKSEQLMNDTKSVLGRITLQQLYEAVASKNNYALYINDGPKPNYSRKY 124
QY 121 GHTKGLLLNVRVQGFLLIHSIPQPIPEBGIDYPTGRRNGSGGICITFKYQYEAIDS 180
Db 125 GHTKGLLLNVRVQGFLLIHSIPQPIPEBGIDYPTGRRNGSGGICITFKYQYEAIDS 184
QY 181 QLLVCNPNVYSCSIPATFHQELIHPQLCTRASSEIPGRLLTTLQSAQGGKFLHFAKSD 240
Db 185 QLLVCNPNVYSCSIPATFHQELIHPQLCTRASSEIPGRLLTTLQSAQGGKFLHFAKSD 244
QY 241 SFDDDDFAAWMAQRLKTHLLTETWORKRQBLPNSCLPYHYVNIKAIKLSRHSYFSSYQD 300
Db 245 SFDDDDFAAWMAQRLKTHLLTETWORKRQBLPNSCLPYHYVNIKAIKLSRHSYFSSYQD 304
QY 301 HAKWCISQKGTNRWTCIGDLNRSPhQAfrSGGfICTQNWQIYQAFQGLVLYYESCK 357
Db 305 HAKWCISQKGTNRWTCIGDLNRSPhQAfrSGGfICTQNWQIYQAFQGLVLYYESCK 361

RESULT 3
```

```
Q9QZK9
ID Q9QZK9 PRELIMINARY; PRT; 356 AA.
AC Q9QZK9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Deoxyribonuclease DLAD.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Wistar;
RC MEDLINE=20025354; PubMed=10558878;
RA Tanuma S.I., ShioKawa D.;
RT "Cloning of a cDNA encoding a rat DNase II-like acid DNase.";
RL Biochem. Biophys. Res. Commun. 265:395-399(1999).
DR EMBL; AF178974; AAP13596.1; -.
DR PIR; JC7131; JC7131.
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR004947; DNase_II.
DR Pfam; PF03265; DNase_II; 1.
SQ SEQUENCE 356 AA; 40472 MW; 5CBC259F42DD6741 CRC64;

Query Match 67.1%; Score 1297.5; DB 11; Length 356;
Best Local Similarity 66.3%; Pred. No. 2e-112;
Matches 236; Conservative 46; Mismatches 73; Indels 1; Gaps 1;

QY 1 MMARLLTSPALLFLGLFGVLGAATISCRNEEGKAVDWFYFKLPKQNKESGETGLEYL 60
Db 1 MTAQPLKAALPLLFLVALSGVLGTPVISINEDGKAVDWFYFKLPRTSRGGTGMGLDYL 60
QY 61 YLDSTTRSWRKSEQLMNDTKSVLGRITLQQLYEAVASKNNYALYINDGPKPNYSRKY 120
Db 61 YLDSTTRSWRKSEQLMNDTKSVLGRITLQQLYEAVASKNNYALYINDGPKPNYSRKY 119
QY 121 GHTKGLLLNVRVQGFLLIHSIPQPIPEBGIDYPTGRRNGSGGICITFKYQYEAIDS 180
Db 120 GHAKGLLLNVRVQGFLLIHSIPQPIPEBGIDYPTGRRNGSGGICITFKYQYEAIDS 179
QY 181 QLLVCNPNVYSCSIPATFHQELIHPQLCTRASSEIPGRLLTTLQSAQGGKFLHFAKSD 240
Db 180 QLLVCNPNVYSCSIPATFHQELIHPQLCTRASSEIPGRLLTTLQSAQGGKFLHFAKSD 239
QY 241 SFDDDDFAAWMAQRLKTHLLTETWORKRQBLPNSCLPYHYVNIKAIKLSRHSYFSSYQD 300
Db 240 FYTDDIDFAAWIAQKLKVLHVESQWRKNHELPSNCLPYHYVNIKAIRGLQSDFFPSHHD 299
QY 301 HAKWCISQKGTNRWTCIGDLNRSPhQAfrSGGfICTQNWQIYQAFQGLVLYYESCK 356
Db 300 HSKWCVSTKDSQARWTCIGDLNRSPhQAfrSGGfICTQNWQIYQAFQGLVLYYESCK 355

RESULT 4
Q9QY48
ID Q9QY48 PRELIMINARY; PRT; 354 AA.
AC Q9QY48;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Deoxyribonuclease DLAD.
GN DNASE2B OR DIAD.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6; TISSUE=Liver;
RC MEDLINE=99428663; PubMed=10497274;
RA ShioKawa D., Tanuma S.I.
```

"DLAD, a novel mammalian divalent cation-independent endonuclease with homology to DNase II.",  
Nucleic Acids Res. 27:4083-4089(1999).  
[2]  
SEQUENCE FROM N.A.  
STRAIN=C57BL/6;  
MEDLINE=21556924; PubMed=11700027;  
Shiokawa D., Tanuma S.;  
RT Head-to-Head with the Genes for Urate Oxidase."  
Biochem. Biophys. Res. Commun. 288:1119-1128(2001).  
EMBL; AF128888; AAF05082.1; -;  
DR EMBL; AF334608; AAL34450.1; -;  
DR EMBL; AF334603; AAL34450.1; JOINED.  
DR EMBL; AF334604; AAL34450.1; JOINED.  
DR EMBL; AF334605; AAL34450.1; JOINED.  
DR EMBL; AF334606; AAL34450.1; JOINED.  
DR EMBL; AF334607; AAL34450.1; JOINED.  
DR MGD; MGI:1913283; Dnae2b.  
DR GO; GO:0005737; C:cytoplasm; IDA.  
DR GO; GO:0005576; C:extracellular; IDA.  
DR GO; GO:0004520; F:endodeoxyribonuclease activity; IDA.  
DR InterPro; IPR00497; DNase II.  
DR Pfam; PF03265; DNase II; 1.  
DR SQ SEQUENCE 354 AA; 40793 MW; 40A5DD6E9AD278A1 CRC64;  
  
Query Match 66.9%; Score 1293.5; DB 11; Length 354;  
Best Local Similarity 66.4%; Pred. No. 4.7e-112;  
Matches 237; Conservative 45; Mismatches 72; Indels 3; Gaps 2;  
  
QY 1 MWRLRTSFALLFLGLFGLVGAATISCRNEEGKAVDWTFYKLPKRNKESGETGLEYL 60  
Db 1 MTAKPLRTVLSLFFALFSGVLGTPETISCRNEYGEAVDWTFYKLPKRTSKASEEAGLQYL 60  
  
QY 61 YLDSTTRSEKSQLMNDKTSVLGRTLQOLYEAVASKNNFAYLIYNDGPKVNYGRKY 120  
Db 61 YLDSTQTKWKSIVYLINSTSALGRTLQHLDYTHNS-TNDTAYLIYNDGVPGSYNGRQY 119  
  
QY 121 GHTKGLLLNRRVQGFMLIHISIPQPPPIPEGDYDPTGRNGSGGICITFKINQYRIDS 180  
Db 120 GHAKGLLVNRTOGFWLIHVSVPFPV--HGYYPTSGRRYGGGICITFGYQFEIDF 177  
  
QY 181 QLLVCPNVYSCISPAFPHQELIHPOLCTRASSEIPGRLLTTLQSAQGKFLHFAKSD 240  
Db 178 QLLVLQPNVYSCIPSTFHWKLIWPMRCANSSSLKIPVYLAELHSAQGLNFVHFAKSS 237  
  
QY 241 SFLEDDTFAAMWAQRLKTHLITETWORKRQELPNCSLPIYHYNKALKLSRHYSFSSYQD 300  
Db 238 FYTDDFTGTMIAQKLKTHLAAQWQKKQELPNCSLPIYHYNKISIGVTSKYSFSSRQD 297  
  
QY 301 HAKWCISQGTKNRWTICIGLNRSPHAPRSGGFICTONWQIYQAFQGLVLYESCK 357  
Db 298 HSKWCVSIKGSANFTWICIGLNRSLHQALRGGFICTKNHYIYQAFHKLVLRYGFCCK 354  
  
RESULT 5  
Q8C589 PRELIMINARY; PRT; 354 AA.  
AC Q8C589;  
DT 01-VAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Deoxyribonuclease II beta.  
GN DNASE2B.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
[1]  
SEQUENCE FROM N.A.  
STRAIN=C57BL/6J; TISSUE=Urinary bladder;  
MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium,  
RX

RA	the RIKEN Genome Exploration Research Group Phase I & II Team;
RT	"Analysis of the mouse transcriptome based on functional annotation of
RL	60,770 full-length cDNAs.";
DR	Nature 420:563-573(2002).
DR	EMBL; AK079224; BAC37579.1; --.
DR	MGB; MG1:1913283; DnaSe2b.
DR	GO; GO:0005737; C:cytoplasm; IDA.
DR	GO; GO:0005576; C:extracellular; IDA.
DR	GO; GO:0004520; F:endodeoxyribonuclease activity; IDA.
DR	InterPro; IPR004947; DNaase_II.
DR	Pfam; PF03265; Dnaase II; 1.
SQ	SEQUENCE 354 AA; 40765 MW; 40AFFB179AD5D428 CRC64;
	Query Match 66.8%; Score 1290.5; DB 11; Length 354;
	Best Local Similarity 66.1%; Pred.No. 8.9e-112;
	Matches 236; Conservative 46; Mismatches 72; Indels 3; Gaps 2;
Qy	1 MMARLLFTSALLFLGLFVGLGVAATISCRNEEGKAVDWFTFYKLPKRKONKESGETGLEYL 60
Dd	1 MTAKPLRTVLSLFFALSGVLGPETISCRNEYGEVDWFIYKLPKRTSKASEAGLQYL 60
Qy	61 YLDSTRWRKSQOLMMDTKSVIGRTLQQIYEAYASKNNNTAYLIYNDGVKPYNYSRKY 120
Dd	61 YLDSTRQTWNKSYLIINSTRSALGKTQLHYDVTHNS-TNDTAYLIYNDGVPGSVNYSRQY 119
Qy	121 GHYKGILLNVRVGFWLIHSIPOPPPPIPEGDYDPPTGRNGSGGICITFKYNQYEAIDS 180
Dd	120 GHAKGLLVNRTOGFWLIHSPVKFPVPV-HGYEYPTSGRRYGQTGITTFGYSQFEEDIF 177
Qy	181 QLLVCNPNVYSCSIPTAFHQELHMPQLCYRASSETIPGRLLITLSAQOGKFHFAXSD 240
Dd	178 QLLVLQNIYSCSIPTSFHWKLIYMRMCANSCLKIPVYLAELHSAOGLNFVHFAXSS 237
Qy	241 SFLDDIPFAAWAQRKLTHLTETWORKROELPNSCLSPYHVINKAIKLSRHVSVFSSVOD 300
Dd	238 FYTDDELFTGWIAQKLKTHLLAQTWQKKQLPNCSLPYHVYNIKISLGTVSKSYFSRRQD 297
Qy	301 HAKWCISQKGTKNRWTCIGDLNRSPhQAFRSGGFICTQNMQIYQAFQGLVLYYESCK 357
Dd	298 HSKWCVSIRGSANRWTCIGDLNLSHLQALEGGGFICTKHVIYQAFHKLRYRGFCCK 354
RESULT 6	
Q8WZ80	
ID	Q8WZ80 PRELIMINARY; PRT; 153 AA.
AC	Q8WZ80;
DT	01-MAR-2002 (TrEMBLrel. 20, Created)
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE	Iung-specific DLAD.
OS	Homo sapiens (Human)
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

RESULT	6
Q8W8ZO	
ID	Q8W8ZO
AC	AC
DT	01-
DT	01-
DT	01-
DE	Lun
OS	Hom
OC	Euk
OC	Mam
CX	NCB
RN	[1]
RP	SEQ
RX	MED
RA	Shi
RT	"Is
RL	Hea
RL	Bio
DR	EMB
DR	GO;
DR	GO;
DR	Int
DR	Pfa
SQ	SEQ
	Query N
	Best L
	Matches
Qy	

Db 1 MPOLCTRASSELIPGRLTTTLOAQOQKFLHFAKSDSLDDIPAAWAOBKLTHTLTETW 60  
QY 265 QRKQELPNCSPYHYVNIKAIKLSHRHSYFSSYQDHAKWCISQKGTQNRWTTCIGDLNRS 324  
Db 61 QRKQELPNCSPYHYVNIKAIKLSHRHSYFSSYQDHAKWCISQKGTQNRWTTCIGDLNRS 120  
QY 325 PHQAFSGGFICTQNWQIYQAFQGLVLYYESCK 357  
Db 121 PHQAFSGGFICTQNWQIYQAFQGLVLYYESCK 153

## RESULT 7

O90759 PRELIMINARY; PRT; 375 AA.  
AC O90759;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Deoxyribonuclease II (EC 3.1.22.1).  
GN CBL1/DNASEB1.  
OS Fowlpox virus (FPV).  
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
OC Avipoxvirus.  
OX NCBI\_TaxID=10261;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FP-9;  
RX Laidlaw S.M., Anwar M.A., Thomas W., Green P., Shaw K., Skinner M.A.;  
RL "Fowlpox virus encodes non-essential homologs of a secreted nematode protein.";  
RT PC-1 and an orphan human homolog of a secreted nematode protein.";  
RL J. Virol. 72:6742-6751 (1998).  
DR EMBL; AJ006408; CAA07012.1; -.  
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.  
DR GO; GO:0016787; F:hydrolase activity; IEA.  
DR GO; GO:0006259; P:DNA metabolism; IEA.  
DR InterPro; IPR004947; DNase\_II.  
DR Pfam; PF03265; DNase\_II; 1.  
KW Hydrolase.  
SQ SEQUENCE 375 AA; 43064 MW; 36C641F74BAD25F6 CRC64;  
Query Match 39.4%; Score 762; DB 12; Length 375;  
Best Local Similarity 44.3%; Pred. No. 1.8e-62;  
Matches 148; Conservative 64; Mismatches 110; Indels 12; Gaps 7;  
QY 26 ISCRNEEGKAVDWFTFFYKLPKRNKESGETGLELYLDSTTRSRKSEQLMNDTKSVLGR 85  
Db 47 VNCVNEEGVVDWVFVYKLPKQ--KLGTKGNEVLYIDSNPNKWKRGKVPINSRYSIIGK 104  
QY 86 TLQOLYAYASKNNTAYLYNDGVPKVPYNSRYKVGHTKGLLWN--RVQGFMLHSTPQ 143  
Db 105 TLPIYDLYDSK--YIYIIFNDGIPGSKYSSKVGHTKGMWNSDSVTGFVLIHSVPR 162  
QY 144 FPPIPEGYDYPPTGRNGSGGICITFKY--NOYBAIDSQLLVNPNVYSCSIPATFHQEL 202  
Db 163 FPPSPVLGYNYPYSGYVYGQSMCLINDLYKGLTALDNTLFPVNFVNCVST--NKL 219  
QY 203 IHMPQLCTRASSELIPGRLTTTLOAQOQKFLHFAKSDSLDDIPAAWAOBKLTHTLTETW 262  
Db 220 NNLVHLCDNKYNTTLYKNVSRWMSRKGEKELTFKSKYFRHIMSAGWIGPTLESDDLSE 279  
QY 263 TWQKRQELPNCSPYHYVNIKAIKLSHRHSYFSSYQDHAKWCISQKGTQNRWTTCIGDLN 322  
Db 280 TWQKRQELPNCSPYHYVNIKAIKLSHRHSYFSSYQDHAKWCISQKGTQNRWTTCIGDLN 337  
QY 323 RSPHQAFSGGFICTQNWQIYQAFQGLVLYYESCK 356  
Db 338 RSPHQAFSGGFICTQNWQIYQAFQGLVLYYESCK 371

## RESULT 8

O9QZK8

ID O9QZK8 PRELIMINARY; PRT; 350 AA.  
AC O9QZK8;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Deoxyribonuclease II.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Wistar;  
RX MEDLINE=20025354; PubMed=10558878;  
RA Tanuma S.; Shikawa D.;  
RT "Cloning of a cDNA encoding a rat DNase II-like acid DNase.";  
RL Biochem. Biophys. Res. Commun. 265:395-399 (1999).  
DR EMBL; AF178975; AAF13597.1; -.  
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.  
DR GO; GO:0006259; P:DNA metabolism; IEA.  
DR InterPro; IPR004947; DNase\_II.  
DR Pfam; PF03265; DNase\_II; 1.  
DR SEQUENCE 350 AA; 38175 MW; 6483AF5AC13C5111 CRC64;  
Query Match 29.0%; Score 560; DB 11; Length 350;  
Best Local Similarity 34.6%; Pred. No. 1.1e-43;  
Matches 123; Conservative 62; Mismatches 150; Indels 20; Gaps 9;  
QY 11 ALLFLGLFGVLGAATISCRNEBKAVDWFTFFYKLPKRNKESGE--TGLEYLYLDSTTR 67  
Db 6 SLLLAALLMW-PAEALSCYDGRPDVWFVYKLP--ANSQSGDKPKWGLMYKYMDQNSE 62  
QY 68 SWKSEQLMNDTKSVLGRITLQOLYAYASKNNTAYLYNDGVPK--PVNYSRYKVGHTKG 125  
Db 63 GWQGVGHIDSKDGAVALITLQPIQ---KNSSQLAFLLYNDQPKSSAODSSRGHTKG 119  
QY 126 LLLMNRVQGFMLHSTPQFP-PIPEGYDYPPTGRNGSGGICITFKYNOYEAIDSQLV 184  
Db 120 VLLDDQGGFMLVHSVPRFPSPASSGAYSWPNARTYGTLLCVSLPSPFPFGIGKQITY 179  
QY 185 CNPNVYSCSIPATFHQELIHPQLCTRASSELIPGRLTTTLOAQOQKFLHFAKSDSLD 244  
Db 180 TYPILVYDHLKLEGIFAQKLPDLLEVTGKHVLRBPNWNSVILTSRAGTTFOSFAKFGKPD 239  
QY 245 DIFAAWMAQRLKTHLLTETWQRKQELPNCSPYHYVNIKAIKLSHRHS--YFSSYQDHA 302  
Db 240 DLTSGLWLAALGTLNLQVFPNPSGILPNSCSGTHKILDITETGFPSPGPTFNATEDHS 299  
QY 303 KWCISQKGTQNRWTTCIGDLNRSRPHQAFSGGFICTQNWQIYQAFQGLVLYYESCK 357  
Db 300 KWCVAPEGP--WVCVGDNRNKRRETHRGGGLTCTQVPALWKAFFSLV--KACK 348  
RESULT 9  
O9J5H0 PRELIMINARY; PRT; 232 AA.  
ID O9J5H0;  
AC O9J5H0;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE ORF FPV032 DNase II.  
GN FPV032.  
OS Fowlpox virus (FPV).  
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
OC Avipoxvirus.  
OX NCBI\_TaxID=10261;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20193820; PubMed=10729156;  
RA Afonso C.L., Tulman E.R., Iu Z., Zsak L., Kutish G.F., Rock D.L.;  
RT "The genome of fowlpox virus.";  
RL J. Virol. 74:3815-3831 (2000).  
RN [2]

SEQUENCE FROM N.A.  
RA Afonso C.L., Tuitman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;  
AL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.  
DR ENML; AF198100; AAF44376.1; -;  
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.  
DR GO; GO:0006259; P:DNA metabolism; IEA.  
DR InterPro; IPR004947; DNase\_II.  
DR Pfam; PF03265; DNase\_II; I.  
SQ SEQUENCE 232 AA; 26772 MW; 14C05BB3BA79C73A CRC64;

Query Match            26.4%; Score 511; DB 12; Length 232;  
Best Local Similarity   43.3%; Pred No. 2.4e-39;  
Matches 101; Conservative 46; Mismatches 78; Indels 8; Gaps 5;

QY 127 LLWN--RVQGFWLIHSIPQFPPIPEGDYDPTGRRNGSGICITFKY-NOYEAIQSLL 183  
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Db 1 MANNSDSVTGFWLIHVSVPFRPPSPVLGVNYFYSGVVGSMWCINLDYKGLTALDNILP 60

QY 184 VCNENVYSCTPATFHOELIHPOLCFRASSETPGRILLTTLSAQAGOKFLHFASDSL 243  
   |||||:|:|:|:|:|:|:~::~::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Db 61 VNPNVNVCSVT---NKNLNANLHYLCNDKNYTLYIKNVRMWSRKGEKLTFAKSIFYR 117

QY 244 DDIFAAWAQRKLTKHLTETWORERQLPNCSSLPYHVNTKAIKLSRHVSFFSYQHAK 303  
   |:|:|:|:|:|:|:|:|:~::~::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Db 118 HDIMSAWIGPTLESDDLSEUWQRGESINTCSSKIYHNVIKSIINVGTST-FINYDHSK 176

QY 304 WCISQKTKNRWTICGLNRSPhQAfRGFCICTQNWIYAOFGLVLYYESC 356

Db 177 WIVSLYDKKG-WVICGINRSPTQRHGGSYGACTNGYLFLKEITVEIEGC 228

RESULT 10

ID QSVED8 PRELIMINARY; PRT; 366 AA.

AC QSVED8;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE CG7780 protein (GH10876p).

GN DNASII OR CG7780.

OS Drosofila melanogaster (fruit fly)

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Rhydroidea; Drosophilidae; Drosophila.  
RN NCBI\_TaxID=7227;  
[1]

RX STRAIN=Berkely.

RC MEDLINE=2013606; PubMed=10731132;  
RX Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.P., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglou L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Beriman B.P., Bhattacharya D., Bolshakov S.,  
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,  
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Clawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Douc L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Posler C., Gabrieliian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.I., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwa C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulpa D., Lai X.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefia,

```

OS Trichinella spiralis (Trichina worm).
OC Eukaryota; Metazoa; Nematoda; Enoplea; Trichocephalida;
OC Trichinellidae; Trichinella.
OX NCBI_TaxID=6334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISS 534;
RA Liu M., Li C., Garcia-Reyna P., Fu B., Roman T., Boireau P.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF331159; AA16519.1;
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR004947; DNase II.
DR Pfam; PF03265; DNase II; 1.
DR SQ SEQUENCE 338 AA; 37155 MW; 44E0897A32795669 CRC64;

Query Match 10.8%; Score 209; DB 5; Length 338;
Best Local Similarity 21.9%; Pred. No. 5.7e-11;
Matches 76; Conservative 55; Mismatches 176; Indels 40; Gaps 12;

QY 23 AATISCRNEGKAVDWFYKLPKRONKESGETGLVLYLDSTTRSRKSEQLMNDTKSV 82
Db 17 AQVATCKNDANDWYFVYKPPNVLSKTLQSGV-----NPAAASRANINGQ--- 65

QY 83 LGRTLQOLYAYASKNNTAYLYNDGVPK--PWNYSRKYGHTKGLLMMNRV--QGFWLI 138
Db 66 AGHSIIRTWASFVVHRAQINVLAYSDDPPNLPNPKSK---TKGVLLVNNAADEAAWV 122

QY 139 HSIPQFPPIPEGDYPTGRRNGQSGICITFKYNOYEAIDSQLVCNPNVYSCSIPATF 198
Db 123 HTVPENFLAY-LNAYSWPPEASTPKGHMFLCVSNKAHLNSVSKAIRYQEPYVYANNLPAAI 181

QY 199 ---HQLIHPQLCTRASSEIPGRLLTTLQSAQOGKFLHFAKSDSLDIDIPAAWAOQL 255
Db 182 LNQMELFNGLINGIDVRVTSFLAHERPATKSVQAVANIQAFGKHSKSFADMYAIRLNR 241

QY 256 KTHLLTETQKRELPSNCSLPVHVNKAIKL-----SRHSYFSSYQDHAKWC-ISQK 309
Db 242 AASIM--VNSPADARSKSICKGHLKQIKTISQLDGVQVSE-----DSAKWALIDGK 293

QY 310 GTRKNRTWTCIGDLNRSFHOAFRSGGFICTONWQIYQAFQGLVLYYESC 356
Db 294 NT-----VCFITNDYTATEKRTPGAACVLENAGVYNAPRTAALNVEAC 336

RESULT 12
Q86QP5 PRELIMINARY; PRT; 339 AA.
AC Q86QP5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Newborn larvae stage-specific DNase II.
OS Trichinella spiralis (Trichina worm).
OC Eukaryota; Metazoa; Nematoda; Enoplea; Trichocephalida;
OC Trichinellidae; Trichinella.
OX NCBI_TaxID=6334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISS534;
RA Liu M.Y., Fu B.Q., Lu Q., Wu X.P., Li C.Y., Wang F., Meng X.P.,
Yao C.Y., Niu T.X., Boireau P.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY174186; AA017050.1;
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR004947; DNase II.
DR Pfam; PF03265; DNase II; 1.
DR SQ SEQUENCE 339 AA; 37406 MW; 6F318435B11B68A2 CRC64;

Query Match 9.9%; Score 191.5; DB 5; Length 339;
Best Local Similarity 21.7%; Pred. No. 2.4e-09;
Matches 80; Conservative 61; Mismatches 180; Indels 47; Gaps 16;

OS Trichinella spiralis (Trichina worm).
OC Eukaryota; Metazoa; Nematoda; Enoplea; Trichocephalida;
OC Trichinellidae; Trichinella.
OX NCBI_TaxID=6334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISS 534;
RA Liu M., Li C., Garcia-Reyna P., Fu B., Roman T., Boireau P.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF331159; AA16519.1;
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR004947; DNase II.
DR Pfam; PF03265; DNase II; 1.
DR SQ SEQUENCE 338 AA; 37155 MW; 44E0897A32795669 CRC64;

Query Match 10.8%; Score 209; DB 5; Length 338;
Best Local Similarity 21.9%; Pred. No. 5.7e-11;
Matches 76; Conservative 55; Mismatches 176; Indels 40; Gaps 12;

QY 23 AATISCRNEGKAVDWFYKLPKRONKESGETGLVLYLDSTTRSRKSEQLMNDTKSV 82
Db 17 AQVATCKNDANDWYFVYKPPNVLSKTLQSGV-----NPAAASRANINGQ--- 65

QY 83 LGRTLQOLYAYASKNNTAYLYNDGVPK--PWNYSRKYGHTKGLLMMNRV--QGFWLI 138
Db 66 AGHSIIRTWASFVVHRAQINVLAYSDDPPNLPNPKSK---TKGVLLVNNAADEAAWV 122

QY 139 HSIPQFPPIPEGDYPTGRRNGQSGICITFKYNOYEAIDSQLVCNPNVYSCSIPATF 198
Db 123 HTVPENFLAY-LNAYSWPPEASTPKGHMFLCVSNKAHLNSVSKAIRYQEPYVYANNLPAAI 181

QY 199 ---HQLIHPQLCTRASSEIPGRLLTTLQSAQOGKFLHFAKSDSLDIDIPAAWAOQL 255
Db 182 LNQMELFNGLINGIDVRVTSFLAHERPATKSVQAVANIQAFGKHSKSFADMYAIRLNR 241

QY 256 KTHLLTETQKRELPSNCSLPVHVNKAIKL-----SRHSYFSSYQDHAKWC-ISQK 309
Db 242 AASIM--VNSPADARSKSICKGHLKQIKTISQLDGVQVSE-----DSAKWALIDGK 293

QY 310 GTRKNRTWTCIGDLNRSFHOAFRSGGFICTONWQIYQAFQGLVLYYESC 356
Db 294 NT-----VCFITNDYTATEKRTPGAACVLENAGVYNAPRTAALNVEAC 336

RESULT 13
Q86QP4 PRELIMINARY; PRT; 342 AA.
AC Q86QP4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DNase II (Fragment).
OS Trichinella spiralis (Trichina worm).
OC Eukaryota; Metazoa; Nematoda; Enoplea; Trichocephalida;
OC Trichinellidae; Trichinella.
OX NCBI_TaxID=6334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISS534;
RA Liu M.Y., Fu B.Q., Lu Q., Wu X.P., Li C.Y., Wang F., Meng X.P.,
Yao C.Y., Niu T.X., Boireau P.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY174187; AA017051.1;
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR004947; DNase II.
DR Pfam; PF03265; DNase II; 1.
DR NON_TER
DR SQ SEQUENCE 342 AA; 37273 MW; CLCB12088ADEAF72 CRC64;

Query Match 6.6%; Score 127; DB 5; Length 342;
Best Local Similarity 19.9%; Pred. No. 0.0025;
Matches 68; Conservative 64; Mismatches 152; Indels 58; Gaps 16;

QY 36 VDMFTFKLPKRONKESGETGLVLYLDSTTRSRKSEQLMNDTKSVLGRTLQOLYEAYA 95
Db 23 VDMFTFKLP-----GKSSSHYL-IPTAATDWTVAANI-DDAQQPIHSTMOK-YFASG 72

QY 96 SKSNNTAYLYNDGVKPNVSNRYGHTKGLL--WNEVOGFWLTHSPQPPIPEGDY 154
Db 73 TKQNTN--IIAYSNPFPFELPMSPCKGVIMAEKIKGFWLVTAKYFNLA----- 124

QY 155 PPTGRRNGQSGICITFKYNOYEAIDSQLVC-----NPNVYSCSIPATFHQLHMPQ--- 207
Db 125 -----GAVGDLFTHEKTTKEA---AAFLCMSYTDVNLRAVAKIIDEQPIVFAQRSA 174

QY 208 -LCTRA--SSSEIPGRLL-----TTLQSAQOGKFLH-----FAKSD-SFLDDIFA 248

```



```

Db 175 AQATQAFYDSSEIQ-KLVNGLHKYQPTASTSGDGIKTLTPPGTVKIFASAPVAYSDDIYS 233
QY 249 ANWAQRKTHLLTETWQRKQELPSNCSLPYHYVNIKAIKLSRHSYFSSVQDHAKWCISQ 308
Db 234 NYVVKLQKSMQVYTGTTATVLRKSCAGTLKVENLGVITVTKDTIPIGQDSARWSVPK 293
QY 309 KGTNRWTCIGDLNRSRPHQAFRSGGFICTQNWQIYQAFQGLV 350
Db 294 --SDSDFVCLNTRGRTANDAKYGASVACVLSKEAALFRKMI 333

RESULT 14
ID Q27073 PRELIMINARY; PRT; 344 AA.
AC Q27073;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 43 kDa secreted glycoprotein precursor.
OS Trichinella spiralis (Trichina worm).
OC Eukaryota; Metazoa; Nematoda; Enoplea; Trichocephalida;
OC Trichinellidae; Trichinella.
OX NCBI_TaxID=6334;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92406752; PubMed=1382055;
RA Vassilatis D.K., Despoimier D., Masek D.E., Polvere R.I., Gold A.M.,
RA van der Ploeg L.H.T.;
RT "Analysis of a 43-kDa glycoprotein from the intracellular parasitic
RT nematode trichinella spiralis.";
RL J. Biol. Chem. 267:18459-18465 (1992).
DR EMBL; M95499; AAA30327.1; -.
DR PIR; A44164; A44164.
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR004947; DNase_II.
DR Pfam; PF03265; DNase_II; 1.
KW Signal.
FT CHAIN
FT SIGNAL
FT CHAIN
FT CHAIN
SQ SEQUENCE 344 AA; 37700 MW; 741D6BE0CF56C28C CRC64;

Query Match
Best Local Similarity 20.7%; Pred. No. 0.022;
Matches 74; Conservative 61; Mismatches 142; Indels 80; Gaps 20;

QY 13 LFLGLFGV-----LGAATISCRNEGKAVDWFTFYKLPKRONKESGETG-LEYLYLDS 64
Db 5 IFLSAFWILNCLQIHAANCTRTATDD-TWFLFLFK-----PVGLLKAKIISP 53
QY 65 TTRWRKSEQLMN-DTKSVLGRTLQQLYEAVASKNNNTAYLLYNDGVKPVNYSRKYGHT 123
Db 54 ANAGWANDGANNTDSGHAL---VQTLAEMWGPILDDMTALGYNTPPKSTITSQTS-S 109
QY 124 KGLLWNR--VOGFWLIHSIPQ-FPPIPEGYDYPPTGRRNGQSGICITFYKNQYEIDS 180
Db 110 KGLMFGNETTDGFWLLHTFERAF--NSVAWSFKFTSEGHMALCLISSEDNVPLI-- 165
QY 181 QLLVCNPNVSCSPATFHQB-LIHMPLQCTRAS-----SSEIPGRLL----- 221
Db 166 -----VPALQYQEWIYFGVSSEKATEFADLTSLIDGSLPTITPPLMNQOT 212
QY 222 LTTLSAQGQKFLHFAKSDSLDIFAAWMAQRKTHLLTETWQRKQELPSNCSLPYHV 281
Db 213 ITTLNSALST--VVYSKSSRLRMEYSGFLAKVMVNM--RIWAVTNTLTQTCG--GKI 266
QY 282 YNIKAIKLSRHSYFSSVQDHAKWCISQKGTNRWTCIGDLNRSRPHQAFRSGGFICTQ 338
Db 267 GFVKVWK-SPVTIDGTQNDRSK-----DKSQNAVIDD---KPVFCFTTNGYSTKQ 312

```

RESULT 15

```

Q27076
ID Q27076 PRELIMINARY; PRT; 315 AA.
AC Q27076;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE P49 antigen (Fragment).
GN P49.
OS Trichinella spiralis (Trichina worm).
OC Eukaryota; Metazoa; Nematoda; Enoplea; Trichocephalida;
OC Trichinellidae; Trichinella.
OX NCBI_TaxID=6334;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91246268; PubMed=2038363;
RA Su X., Prestwood A.K., McGraw R.A.;
RT "Cloning and expression of complementary DNA encoding an antigen of
RT Trichinella spiralis.";
RL Mol. Biochem. Parasitol. 45:331-336 (1991).
DR EMBL; M64242; AAA30328.1; -.
DR GO; GO:0004531; F:deoxyribonuclease II activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR004947; DNase_II.
DR Pfam; PF03265; DNase_II; 1.
FT NON_TER
FT CHAIN
SQ SEQUENCE 315 AA; 34279 MW; 9AAE2AC93A04CCDE CRC64;

Query Match
Best Local Similarity 20.2%; Pred. No. 0.32;
Matches 66; Conservative 58; Mismatches 130; Indels 72; Gaps 18;

QY 37 DWFTFYKLPKRONKESGETG-LEYLYLDSSTTRSWRKSEQLMN-DTKSVLGRTLQQLYEAY 94
Db 6 EMFLFLFK-----PVGLLKAKIISPANAGWANDGANNTDSGHAL---VQTLAEMW 52
QY 95 ASKSNNTAYLLYNDGVKPVNYSRKYGHTKGLLWNR--VOGFWLIHSIPQ-FPPIPEG 151
Db 53 GPILDDMTALGYNTPPKSTITSQTS--SKGLMFGNETTDGFWLLHTFERAF--NSVA 109
QY 152 YDYPPTGRRNGQSGICITFYKNQYEIDSQLLVCNPNVSCSPATFHQB-LIHMPLQCT 210
Db 110 WSWFSKFTSEGHMALCLISSEDNVPLI-----VPALQYQEWIYFGQVSS 154
QY 211 RAS-----SSEIPGRLL-----LTTLSAQGQKFLHFAKSDSLDIFAAWMA 252
Db 155 EKATEFADLTSLIDGSLPTITPPLMNQOTITTLNSALST--VVYSKSSRLRMEYSGFLA 212
QY 253 QRLKTHLLTETWQRKQELPSNCSLPYHYVNIKAIKLSRHSYFSSVQDHAKWCISQKTK 312
Db 213 KVMVNM--RIWAVTNTLTQTCG--GKIGFVKVWK-SPVTIDGTQNDRSK-----DK 260
QY 313 NRWTCIGDLNRSRPHQAFRSGGFICTQ 338
Db 261 SQNAVIDD---KPVFCFTTNGYSTKQ 283

```

Search completed: October 14, 2004, 00:31:14  
Job time : 58.7384 secs

***This Page Blank (uspto)***